"Real-Time Interactive Secure Forensic System (RTISFS)"

November 29, 2011

Sponsored By

Defense Advanced Research Projects Agency (DOD)
Information Innovation Office

ARPA Order BK20-00

Issued by U.S. Army Aviation and Missile Command Under

Contract No. W31P4Q -11-C-0213

Name of Contractor: Outdo Inc.
Principal Investigator: Frank J. Sauer
Address: 1534 N Stafford Street
Arlington, VA 22207-3108
Phone: 240-476-2732

Effective Date of Contract: April 18, 2011
Short Title of Work: Final Report

Contract Expiration Date: December 21, 2011
Reporting Period: September 30, 2011 to November 29, 2011

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
11/20/2011	Final Report	SEP 30 - NOV 29, 2011

4. TITLE AND SUBTITLE
Real-Time Interactive Secure Forensic System (RTISFS)

5a. CONTRACT NUMBER
W31P4Q-11-C-0213

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Sauer, Frank, J; Lukasik, Stephen, J, Ph.D.; Yengst, William, C; Fritz, W, Ermarth; Townes, Miles, D; Givner-Forbes, Rebecca; Hunter, G; Kenneth; Russell, Ted; Pollard, Neal; Smith, Sharon, S, Ph.D.; Lipson, Randall, H; Charney, David, L, M.D.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Outdo Inc. 1534 N Stafford Street
Arlington, VA 22207-3108

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Defense Advanced Research Projects Agency
Information Innovation Office
3701 North Fairfax Drive
Arlington, VA 22203-1714

10. SPONSOR/MONITOR'S ACRONYM(S)
DARPA; I2O

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

RTISFS will support identifying and defending against malicious insiders functionalities through: a wider range of access limitations; dynamic environment supporting interactions with users without revealing the depth of forensic and enforcement capabilities; scripted interrogatories to assist separating anomalies attributed to malicious insiders from those of honest intent;ability to increase levels of surveillance or limitation of access as increasing suspicion dictates to minimize damage; and extendable scripting language for handling various types of anomalies tailored for the subject domain. RTISFS will accomplish this according to all applicable legal procedures in such a way that all potential response options are maintained: legal action, turning,use of insider as unwitting communication channel, and collection and penetration of the adversary actor. The proposed system will support a much greater range of and finer degrees of access control.

15. SUBJECT TERMS
Computer security, insider, counter-intelligence, fraud, forensics, law enforcement, surveillance, cyber espionage

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Frank J Sauer
U	U	U	SAR	100	19b. TELEPHONE NUMBER *(Include area code)* 240-476-2732

Reset

"Real-Time Interactive Secure Forensic System (RTISFS)"

November 29, 2011

Sponsored By

Defense Advanced Research Projects Agency (DOD)
Information Innovation Office

ARPA Order BK20-00

Issued by U.S. Army Aviation and Missile Command Under

Contract No. W31P4Q -11-C-0213

Name of Contractor: Outdo Inc.
Principal Investigator: Frank J. Sauer
Address: 1534 N Stafford Street
Arlington, VA 22207-3108
Phone: 240-476-2732

Effective Date of Contract: April 18, 2011
Short Title of Work: Final Report

Contract Expiration Date: December 21, 2011
Reporting Period: September 30, 2011 to November 29, 2011

Distribution:

4 copies (electronic only)

Director
Defense Advanced Research Projects Agency
ATTN: I2O (Dr. Rand Waltzman, Program Manager, rand.waltzman@darpa.mil)
3701 North Fairfax Drive
Arlington, VA 22203-1714

Director
Defense Advanced Research Projects Agency
ATTN: I2O (Wendy M. Smith, ADPM, wendy.smith@darpa.mil)
3701 North Fairfax Drive
Arlington, VA 22203-1714

Thomas G. Bramhall (thomas.g.bramhall@us.army.mil)

DARPA SBIR Office (sbir@darpa.mil)

4 copies (printed)

U.S. Army Research, Development & Engineering Command (1 Copy)
ATTN: RDMR-AS-RSIC (Documents)
Bldg 4484
Redstone Arsenal, AL 35898-5241

Director (1 Copy)
Defense Advanced Research Projects Agency
ATTN: OMO/DARPA Library
3701 North Fairfax Drive
Arlington, VA 22203-1714

Director (2 Copies)
ATTN: Acquisitions/DTIC-OCP, Rm-815
8725 John J. Kingman Rd., STE 0944
Ft. Belvoir, VA 22060-6218

Table of Contents

OVERALL PROJECT CONCLUSIONS
TO BE DRAWN FROM THE SEPARATE BIMONTHLY REPORTS

Stephen J. Lukasik

Unlike the Executive Summary, that treats this final report as the last in a series of three such reports, the purpose of these Conclusions is to cover the *entire* Phase I SBIR contract as a whole.

From the outset we viewed the critical task as establishing *who* we were looking for. Until we could establish this, any effort to fashion a combined online trap and offline analysis regime would be hit or miss.

In this, "who" can obviously not mean a personal identifier, since this would be a circular definition. Granted we are looking for our targets through their online presence, our targets are defined through their commission of predefined suspect actions.

Our software system to do this, RTISFS, was to constitute an early warning system. It would contain flexibly definable sensors, implemented in software, that would record actions corresponding to definable criteria and alert a counterintelligence controller (CIC) of that fact. All RTISFS records would be preserved and protected from tampering, accessible only to a facilities CIC, and searchable with offline analysis tools that would be part of the RTISFS system.

We recognized there were three conditions the system would have to meet to perform its intended early warning function: false alarms had to be minimized; all collection and use of data had to meet constitutional and derivative legal requirements; and the system had to function in the background to preserve all options for dealing with users seen as possibly violating the trust placed in them attendant upon their access to sensitive information. By the nature of our design, the false negative problem did not bulk as large as the false positive problem in our estimation. The users would be with us for a long time and hence we would have multiple shots at detecting their trust violations.

From our work, we draw the following conclusions:

1. There are two places to look for hints about how an inside online user might behave: the past and the future. The past offers advantages: real events, real people, real public records. But it is history. People one really does not know to any depth; circumstances selected for reporting by others; public records that are the result of plea bargaining; need to protect sources and methods of identifying the insider; cover up of prosecutorial errors; political influences; old technology on both sides; motivations no longer relevant. *Caveat emptor*.

The future is speculative. It has not happened yet; it may happen or it may not; technology will change; people will change; motivations will change; laws will change. Politics will change. *Caveat emptor*.

2. Bearing these cautions in mind, we collected information of various types and quality from a variety of sources from the past, and assembled a set of studies looking at the future. These studies had as their objective to identify insider motivations, methods, objectives, the number of them at any time, and means of identifying them through their online behavior based on unconscious clues in their use of language. On balance, we find the information we have collected to be a useful base for projecting future actions and for suggesting implementable ways to operate an RTISFS early warning system.

3. Lacking alternatives to inevitably flawed data, we have used them to the extent possible always keeping their limitations in mind. Dividing the data into Cold War (1949–1989) and post-Cold War (1999–2010) we see very different insider behavior. Much of our intuition derives from our long Cold War experience, and for trained agents of major foreign powers, we find it as valuable as expected.

4. We find post-Cold War insider behaviors to derive from very different motivations, and, we believe, in attitudes toward information influenced by the information age introduced by the WorldWideWeb and the attitudes of young people toward privacy, value of information, and the degree to which information should be shared. It is this post-Cold War perspective that has caused us to draw heavily on future projections as a basis for estimating the 21st century counterintelligence environment.

5. The SBIR program directions caused us some difficulty in choosing between what seemed to be different development goals. We saw the DARPA intent was to national security needs interpreted narrowly: government employees and government contractors dealing with sensitive national security information and keeping it from foreign entities intent on attacking U.S. security. To deal with such national security threats, a counterintelligence controller has a reasonably broad mandate in examining user behavior. This is, if you will, the "easiest" case where the insider has the fewest rights to privacy and the investigator has the greatest degree of flexibility. This has, therefore, been our primary focus of attention in defining RTISFS functionality.

6. But the SBIR program envisions a commercial product, calling, for example for a marketing plan, to fulfill its mandate to assist small businesses develop innovation products of relevance. The focus of the commercial market is not only to protect classified and proprietary information, but to a much greater extent to protect against financial fraud. The organizations benefitting from such a product are largely commercial, and while having civil and criminal statutes and standards to protect them, they also operate in a world where the overriding flexibilities of national security in employee supervision are unavailable. Civil rights, commercial and employee contracts, regulations, and the burdens of proof to establish liability and guilt are very different in the commercial world. We have, therefore, chosen for the present to defer addressing these circumstances in our software sensors, our diagnostic tools to detect deception, and in detailed understanding of the legal framework within which our detection principles would operate.

7. With respect to the applicability of our detection and assumed investigation processes, we found that the current legal framework for the protection of sensitive national security information interposed no unusual constraints. This is the result of our security architecture

where the RTISFS software was restricted to the role noted above: early warning of evidence of deceptive behavior in already vetted federal employees and contractors. We have designed RTISFS so that it is never called upon to produce evidence intended to be used in a judicial proceeding. That way there is some expectation that its processes and algorithms can remain secure from being gamed or avoidance behavior learned. RTISFS is an aid in the discovery of leads and clues, each then followed up in accordance with current constitutional protections and due process under existing statutes.

8. To an extent, therefore, RTISFS must contend with the same needle-in-a-haystack problem counterintelligence currently faces. We have addressed the problem of a large search space in three ways. First we limit the search space by limiting the data to be examined to that most likely to be relevant to aberrant behavior, regardless of its potential seriousness. This is done through the definition of the online behavioral sensors to those most likely related to security-related trust violations.

9. The second way to limit the search space is to record potential violations and generate statistical measures dealing with potential seriousness, number of such violations, the rate of such violations, and changes in user behavior from baseline behavior to atypical behavior. In this way, at any time the CIC has available an ordered list of users ranked by their apparent security threat.

10. The third approach we take is to construct interrogatories to inquire as to selected user behaviors. These interrogatories are formulated to elicit free-form user textual input. This input is analyzed offline using new results from the field of psycholinguists, where differences in word use, sentence structure, and timing can be examined. RTISFS would be accompanied by a variety of offline analysis tools drawn from an evolving professional literature. Through this three step, largely automated approach, the CIC can establish a threat metric for users to guide the allocation of offline manual criminal investigation.

11. Critical to this semi-automated process is the establishment of non-threatening baselines for users, from which aberrant behavior can be recognized. We have found that this can be done through examination of trust violations from historical data. Based on a wide range of circumstances, we find that in a homogeneous occupation population, the incidence rate of trust violations, when the probability of being identified or when the penalty for being caught is relatively small, is 10^{-3}, one in a thousand.

12. We find when the stakes to the violator are higher, the incidence drops to around 10^{-4}, one in ten thousand. Thus in a large user organization, one can expect that 99.9% of users will not violate the trust placed in them. Therefore "honest" baselines can be established on the basis of date collected from samples of users that are still reasonably large.

13. We find the number of software sensors, for practical application, should be ≤ 5, and the number of query response interactions with a user should be ≤ 3 when RTISFS is in its semi-automated collection–discovery mode.

14. We find that Cold War lessons learned from dealing with the USSR/KGB/GRU security apparatus must be reevaluated since major state adversaries can be expected to change their operational procedures in response both to the new availability of web-based information as well as to respond to specific information needs. We have identified Russia and China as the two most important states challenging the U.S. Both will use easily available web information to minimize the use of more limited and less immediately responsive insider collection opportunities. Military technology and military threat information will continue to be as important as during the Cold War. Adding to this, however, is the increasing importance of industrial information collection far beyond purely military needs.

15. The most aggressive collector of sensitive U.S. information is China. Their collection process is to assemble large amounts of information, mainly open source but with targeted insider probes, and to be guided by what the mosaic of information suggests. The Chinese intelligence operation also includes a large number of agents engaged in influence operations. These are largely not insiders and thus their efforts will be evident only through the changes in the online behavior of insiders they induce.

16. The second aggressive collection activity targeted against the U.S. is that of Russia. It is here that the greatest utility of our Cold War analysis is to be found. Like the Chinese, the Russians also take a long-term approach to intelligence. This takes two forms. One is also influence operations where the intent is not to find out what the U.S. will do, but the *shape* what the U.S. does. Like the Chinese, Russians also favor the use of long-term "illegals." These are agents intended to work their way into U.S. society and culture, thereby developing a strong record of trust and apparent loyalty to the U.S.

17. We believe, however, that future intelligence collection will be dominated by cyber penetration of information systems to supplement open source and personal insider collection. To this end, we believe the insider role will be the opposite of find-and-remove. The insider will not be an information collector but will be a route for inserting trapdoor software into closed systems so the lengthy and detailed search and removal function can be done more easily and efficiently by large numbers of search/collection/analysis workers in outside locations.

18. We are deeply suspicious of the ability of R&D groups to avoid being seduced by the beauty of their proposals. Therefore we have subjected our ideas to a Red Team analysis of our efforts. We have identified seven RTISFS circumvention routes. This does not, at this point in its design, say the system is without value. It points the development team to areas where RTISFS must be strengthened before deployment. It also points the deploying organization to those threats it must address separately because they lie beyond the scope of threats RTISFS can be expected to protect against.

19. As we visualize the future counterintelligence environment, we are taken by the view that while counterintelligence agencies can usefully improve their processes, much more important is its concepts of targets. The most severe threats will continue to assault us, but the volume of sensitive information outflow will increase by orders of magnitude. This will derive from increasing transparency in the U.S., only part of which will be by public policy, but also by changing views of what should remain private, and by loosening loyalty to sovereign states.

Globalization, divided loyalties, voluntary information sharing, and the natural utility of social media enhanced affinity groups will shift the balance between information that is common and that which is not.

20. As we have undertaken research to sharpen our proposed idea for the ADAMS SBIR program, we have arrived at two positions quite different from those with which we started. First we see important ideas for more fundamental research on the mental processes that lie behind perceptions of truth and the importance of truth as a motivator of behavior. The second is we see how the details of software mechanisms for accomplishing the goal of the ADAMS program are relatively unimportant compared to the former. As we came to this recognition we reduced our efforts on Task 3 and transferred research resources to the examination of the more fundamental issues addressed in this report.

21. That said, we conclude that there is a fundamental flaw in the formulation of the ADAMS program. We see the collection of masses of largely irrelevant data, coupled with clever tree-pruning algorithms and massive computing power to be self-limiting. If technology capabilities are invested in the collection of increasing amounts of irrelevant data, the signal-to-noise ratio in the search space diminishes rather than increases. The critical technical frontier rests with understanding deception at its deepest level, not with energetic manipulation of digitally-encoded information.

EXECUTIVE SUMMARY
FOR FINAL REPORT

We have interpreted "final report" as the last report of a set of three, each covering a two month period of work. Thus, the final report is the sum of all three. But to reflect the end of the contract, this report (#3) has, in addition to this Executive Summary, an overall set of conclusions based on all the research that has been performed to date.

TASK 1 CREATE A TAXONOMY OF INSIDER MODELS FUTURE PROJECTS WILL HAVE TO ADDRESS

In this final period we continued to examine the incidence of trust violations because it is important to be able to establish user baseline behavior of trustworthy people so that the deviations for untrustworthy people can be reliably established. In the previous period we examined cases of espionage over a long period of time. We found the incidence of trust violations varied over several orders of magnitude, generally $\leq 10^{-4}$. To test this incidence rate on a larger and different population than espionage, we looked at other cases: Federal judiciary, customs and border control, transportation security inspectors, terrorism, and violent crime. The following Table 1 summarizes the complete set of population incidence rates.

Table 1
Incidence of trust violations

MAJOR CATEGORY	SUBCATEGORY	DECADE(S)	NO. OF CASES	INCIDENCE RATE	PERCEIVED SERIOUSNESS
Espionage	U.S. Revolution	1770	10	7×10^{-4}	Not
	U.S. Civil War: Union	1860	3	1×10^{-4}	
	U.S. Civil War: Confederate	1860	3	5×10^{-7}	
	WW II Venona +	1940	600	2×10^{-4}	
	Cold War	1950-2000	44	1×10^{-5}	
Military combat	Alamo: Texas	1830	1	4×10^{-3}	
	Alamo: Mexico	1830	1	4×10^{-4}	
	Custer's Last Stand	1870	2	6×10^{-3}	Not
	Rorke' Ridge (Zulu War)	1870	1	6×10^{-4}	
Federal Judiciary	U.S. Civil War	1860	1	2×10^{-3}	Not
Homeland Security	Border Control	2000	129	$1\text{–}2 \times 10^{-3}$	Not
	Transportation Security	2000	488	1×10^{-3}	Not
Terrorism	Worldwide	2000	>10,000	10^{-5}–10^{-2}	
	U.S.	2000	8,605	$<10^{-5}$	
Violent Crime	Worldwide (incl. U.S.)	2000	Unk.	10^{-3}–10^{-2}	Not

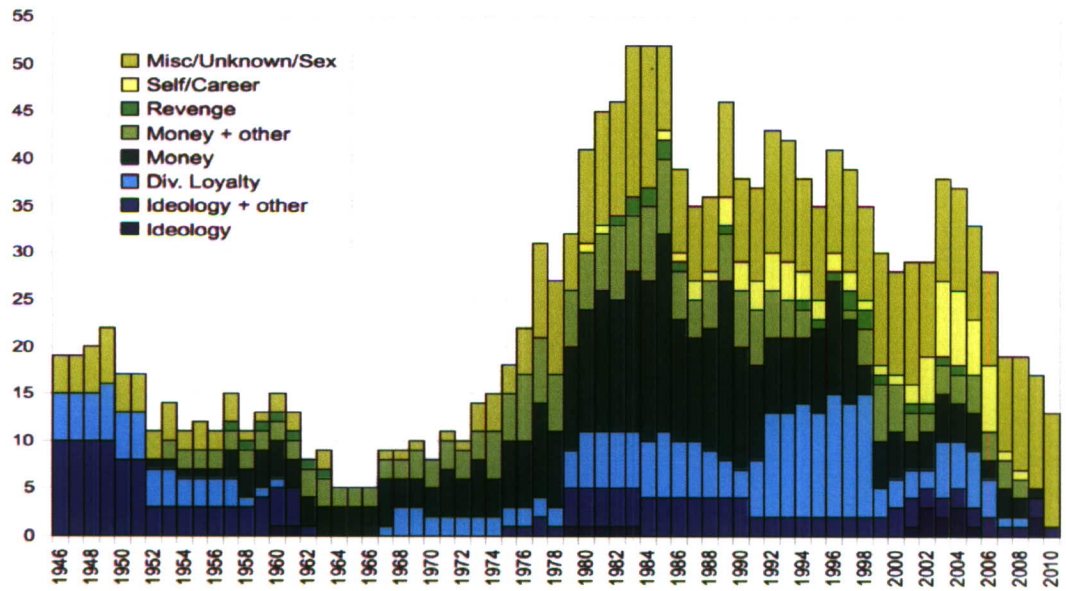
We find that in cases where the penalty for the violation of trust is not seen as serious, the incidence rate goes up from the 10^{-4} level to a higher level such as 10^{-3}. So our conclusion that

establishing stable baselines will not be unreasonably difficult continues to appear justified as long as the size of the learning set is smaller than 1/incidence rate.

We then shifted from using historical data to size the problem to considering the future environment for counterintelligence. This is summarized in Figure 1 below. We have seen in our data that post-Cold War the reasons for trust violations are significantly more varied. The data show in the 80's and 90's a growth of cases where money and divided loyalty became increasingly common motivations. Post-2000, while money and divided loyalty remain important, self/career and miscellaneous reasons become more important.

The second most important intelligence effort directed against the U.S. is that of Russia. Our Cold War cases presented a great deal of information on a range of Soviet espionage, but they did not present the full scope of the Soviet, and current Russian, repetoire. The analysis presented here updates Russian intelligence objectives in economic areas, both for the State as well for the intelligence apparatus itself. It addresses long-term Russian invesments in illegals, illustrated by the group exposed in 2010. But we suggest that U.S. intelligence "success" may be have been of Russian deception operation to cause us to believe we had "fixed" the problem when in fact they were decoys to accomplish precisely that. This illustrates a third Russian intelligence goal, not of learning what the U.S. thinks, but to *influence* what the U.S. thinks. This simplifies some intelligence collection since the Russians, if successful, have a good idea already of where they led us. Such activities create the need for a different kind of insider, not to remove something but simply to be able to report how much of their influence operations were successful.

Figure 1
Espionage activity by motive, 1946 – 2010



We explored whether globalization could be important. To this end we examined situations where these factors were relevant: naturalized citizens, dual citizens, and diminished citizenship, meaning where even for native U.S. citizens their tie to government is seen as weakening. We

find these factors to be increasingly relevant. We also examined the case of disloyalty by contractors where the government's presence is less apparent in the workplace. The picture that emerges is that confidence in government is weakening and, with it, loyalty is seen more flexibly. We do present statistical data that shows the frequency of disloyal insiders is significantly higher among contractors. Granting the U.S. can not change, nor does it want to change the trend to globalization, these data offer opportunities to "tune" RTISFS to refine searches for trust violations.

Finally we examined the case of leaking secrets. While the phenomenon is by no means new, prosecutions for leaking are increasing, though successful prosecutions are problematic at this point. Leakers frequently have genuine feelings of loyalty and can be seen more as cases of divided loyalty than of disloyalty. For present purposes, we simply note that Cold War espionage cases do not provide useful indicators for RTISFS to defeat this set of base-level unauthorized disclosures.

TASK 2 RESPONSES TO CHANGES IN USER BEHAVIOR ONLINE

In Report #2 the basic design of the interrogatory process that is central to eliciting evidence that could indicate deceit by users was outlined. It pointed to the potential logical complexity of the RTISFS query–response system and the need to keep the number of trigger events and user–system interrogation interactions small.

This paper extends the design of the interrogatory process by going to the next-level down, and examining how interrogatory questions might be framed and what linguistic analysis metrics can be used to base assessments of possible deceit. Linguistic analysis is seen as a two-level process: a set of tools that can be used online to guide automated RTISFS responses and tools only useful offline for more in-depth and longer term assessment of users as potential foci of traditional investigatory processes.

Specific guidance is offered on relating preprogrammed trigger sensors to consequent interrogatories through the combined use of open-ended and "bait" questions.

Since the RTISFS approach of using unconscious linguistic cues indicative of a users mental state rests on recent research on psycholinguistics, it is limited by the current uncertainties in that developing science. We examine one class of user whose inherent view of truth and falsity differs from that of the general population: psychopaths and sociopaths. In view of possibility that such individuals may constitute 4% of the population, their isolation from the truth could set a base level for the ability of RTISFS to meet the intended program goals.

Ways of separating these users from the remaining 96% of a typical user population are suggested. What this says, more generally, is that our taxonomy will have to be expanded to recognize major subsets of psychological behavior. The analysis also indicates how the native language of a user can impact assessments of deceit based on psycholinguistic analysis.

In view of this central dependence of RTISFS on its essential scientific foundation, we outline a set of steps recommended to further explore this approach to identifying potential weak links in the protection of sensitive information.

The following paper continues the analysis of fundamental limits on counterintelligence as currently organized and practiced. RTISFS, and the ADAMS program more generally, forces one to examine not simply current practices but the fundamentally changed nature of protecting information in the 21st century information age. The inherent processes formed during WW II and the Cold War are no longer completely adequate. There is no longer a big difference between insiders and outsiders. The bulk of the foreign agents working against U.S. security are not the few insiders with unique access, but the outsiders with total access, first to public information, and then to public information technology to reach non-public information.

We see counterintelligence not as a largely reactive enterprise using weak tools to vet insiders, but as a proactive enterprise joined in a cooperation with the separate functions of personnel, physical, and communications security. The bureaucratic and political origin of this separation seems obvious. It is an increasingly unsatisfactory separation in the current information environment.

Finally, in the spirit of testing our RTISFS construct to identify its weaknesses for remediation, we undertook a small Red Team effort at the end of the program. We find that, subject to the psychological science uncertainties already identified, RTISFS is not a silver bullet. It is part of a set of actions that can increase the protection of some sensitive information.

We conclude that intelligence and counterintelligence must recognize what managers call tradeoffs and what physicists call uncertainty principles:

For the defender: $\Delta(\text{security})\ \Delta(\text{utility}) \geq \text{some constant}$

For the penetrator: $\Delta(\text{precision})\ \Delta(\text{timeliness}) \geq \text{some constant}$.

Furthermore no security system, certainly not RTISFS, can bypass the question of "Who watches the watchers?" Nor can one think that technology alone can substitute for error-prone humans, or detect the errors of human users. The interaction of attackers and defenders is the invariant in human relations. RTISFS can be an important and useful factor in moving that interaction in the direction of increasing security.

TASK 3 DESIGN CRITERIA FOR A FUNCTIONALLY-COMPLETE RTIFSF

Parts of requirements outlined in the Phase I proposal for Task 3, relating to how the proposed interactive defensive system can be protected from gaming by an insider have been described under Task 2 in this report. The answers depend more on the constraints and limits driven by legal and operational requirements than on the software. The operation of RTISFS will be controlled by facility security doctrine and needs. The software design will support a wide range of operational needs.

This is also the case for the issue of constraints on monitoring, recording, and retaining baseline data. For an RTISFS deployed in a classified environment, government or contractor, we found in the Task 2 work that there are none. The amount of data collected by RTISFS is, by its fundamental design principle of minimizing the size of search space, trivial for current computing technology. While we have not constructed a sufficiently detailed design model to definitively answer the performance penalties question, earlier work assessed it to be in the range of 1–2%.

Specification of the necessary software modules is presented here. The amount of time needed to reach TRL 8, software ready to be installed in an operational location, is 48 months from the intended start of a software development effort. The SBIR program speaks of a Phase II as 24 months and anticipates that in Phase III, seen to be another 24 months, that some non-government funding will be available.

As described in this final report, in addition to RTISFS, offline software tools will be required as well. These will be commercially available linguistic analysis tools. We point to the utility of two such software packages but, at this stage, no decision can be made further work. Interfacing such offline tools will require, at best, simple interface code to allow RTISFS and offline tools to exchange data files.

TASK 1 – TAXONOMY OF INSIDER MODELS FUTURE PROJECTS WILL HAVE TO ADDRESS

THE LIKELIHOOD OF VIOLATIONS OF TRUST

William C. Yengst[1] and Stephen J. Lukasik

INTRODUCTION

In designing any system one needs a way to see if it is working, and if it is working, whether it meets expectations. In the case of RTISFS, "working" can not be established by the simple fact that its software does not crash the system in which it is installed. One has to know if it is identifying among the system's users those who may be violating the procedures established to prevent the disclosure of the information contained in the system.

In the case of national security information, its unauthorized disclosure to foreign governments constitutes espionage, although its disclosure may result in judicial punishment for lesser offenses. We examined cases of U.S. espionage over time, from the Revolutionary War to Cold War and the current post-Cold War period. Espionage is an uncommon event. Most people are law abiding, and unauthorized disclose is related to the likelihood that loyal citizens will find themselves in situations, sometimes unwittingly, where they do not adequately discharge the trust placed in them.

This Table 2 summarizes the data we found through examining a range of U.S. historical national security cases:[2]

Table 2
Historical national security cases by time period

TIME PERIOD	NO. OF SPIES DETECTED	INCIDENCE RATE
Revolutionary War	10	7×10^{-4}
Civil War Union side	3	1×10^{-4}
Civil War Confederate side	3	5×10^{-7}
WW II Venona collection	600	2×10^{-4}
Cold War average	44	1×10^{-5}

The incidence rates shown in the last column range between 10^{-6} to 10^{-3}, one in a million people, and one in a thousand. Thus identifying them ranges from detecting something quite rare, and requiring quite sensitive detectors, to something whose probability might be characterized as "small."

Anyone who must measure anything recognizes that the care needed to make a measurement to one ppm is different from one ppt. Masking background noise and the impact of other naturally

---

[1] Mr. Yengst passed away before the completion of this paper.

[2] "Real-Time Interactive Secure Forensic System (RTISFS)," submitted to the Defense Advanced Research Projects Agency, 1 October 2011.

occurring effects have to be understood if one is not to be misled. In signal processing terms, one needs a signal-to-noise effectiveness of at least 30 to 60 db to find cases of espionage.

In the narrowest context, the above is adequate for addressing the identification of cases of espionage. But our effort to identify unauthorized disclosures of national security information is only part of a larger growing problem of violations of trust. National security violations are one area of applicability of RTISFS. But violations of financial trust are important in many other situations. Aside from financial losses, one often finds that theft or money or materiel are used to fund other types of illegal activity. We also find, in the area of information security, that motivations of many post-Cold War national security violations can be quite apart from a desire to harm the national security of the U.S. These arise in part from changing attitudes of people regarding information and their national security obligations. In even more cases, particularly young people, they do not see obligations of even personal privacy as important, and the plethora of so much readily available information debases the value of much of the rest. Most of it is free and, by extension, *all of it* should be free.

The examination of the incidence of trust violations beyond our original scope has been undertaken because we recognize that, not only RTISFS has broader areas of potential applicability, but that future attitudes regarding the importance of protecting important information appear to be changing as more as free-information and as universal access to information changes norms of behavior.

What follows, therefore is an examination of other classes of trust violations. These are: violations of expected behavior in combat; failure to respect oaths take by judges to uphold the Constitution; corruption of Border Control and Customs inspectors; theft by inspectors screening baggage passengers on airlines; commission of violent crime; recent FBI counter-terrorism initiatives; the incidence of international armed terrorists; and computer crime. A number of these areas have been heavily researched. We have by no means been exhaustive, for that would go far beyond our original intent. We have tended to historical situations because we are interested in the beginnings of things.

What has driven our direction is the simple question of incidence rates, so that we can design sensors fully aware of the degree to which our sensors will require signal processing gain.

TRUST VIOLATION IN MILITARY COMBAT

Battle of the Alamo – During the Texas Revolution, Mexican troops under President-General Antonio Lopez de Santa Anna with an army of 2,400 troops laid siege to the Alamo Mission near San Antonio. The mission was defended by approximately 260 soldiers (100 Texan militia and reinforcements led by co-commanders James Bowie and William B. Travis.) The siege started on 23 February 1836, when Santa Anna marched 1,500 troops into San Antonio de Bexar and declared that he had come to reclaim Texas. For the next 12 days, the two forces engaged in several skirmishes, each driven back by accurate artillery and rifle fire from the defenders. Aware that the garrison could not withstand a large scale assault, Colonel Travis wrote multiple letters asking for more men and supplies. Fewer than 100 reinforcements arrived during the siege.

Early on 6 March, the Mexican army attacked in mass. The Texans repulsed two attacks but were unable to prevent the walls from being scaled. The defenders withdrew to interior buildings of the mission. Those who attempted to surrender or flee were tracked down by Mexican cavalry. Between five and seven Texans surrendered and were executed. Eyewitness accounts of the battle reported up to 257 Texans were killed. The Mexicans lost 400-600 killed or wounded. All but two defenders were killed.[3] The Mexicans released two dozen women and children captured in the mission plus Bowie's slave Sam and Travis's slave Joe.

This account yields two data points. Henry Womell escaped but died of his wounds three months later.[4] Brigido Guerrero, a Mexican army deserter who had joined the Texans, was spared by the Mexican commander. One Texan out of 284 for an incidence of 4×10^{-3} and one Mexican out of 2,400 for an incidence of 4×10^{-4}.

Custer's Last Stand – Under General Alfred Terry's command with 2,500 cavalry and infantry troops in North Dakota, General George Custer with 647 troops (eight companies of cavalry) set out in early June 1876 to find the Sioux Indians under leadership of Chief Sitting Bull. Their camp stretched for six miles and its thousands of horses had eaten all the grass and had to move. As he approached the Little Big Horn valley, Custer separated his forces, keeping five companies, and giving one each to Captain McDougal for rear guard, Major Reno to scout, and Captain Frederick Benteen to prevent the Indian village from escaping through the upper valley. Sitting Bull had 6,000 warriors. His total force was outnumbered by at least three to one.

When Reno spotted the Sioux village near Rosebud River on 25 June, Custer ordered Reno to attack the village while he sent for the remainder of his force to join and bring up their two cannons. Reno with 175 men, swept through the village, killing many, but realized his force could be trapped and ordered a retreat. Within minutes the Indians found Custer with only 210 men and forced them back to a ridge to the north. As the Indians closed in, Custer ordered his troops to kill their horses, stack the carcasses, and use saddles to form a wall for protection.

The battle along the ridge lasted all day during which Reno's and Benteen's forces united. When the Indians broke off fighting for the night, Reno's and Benteen's surviving troops, including wounded, escaped. At dawn the next morning, a desperate battle ensued until about 10 AM, with three or four massive Indian attacks on Custer's positions.[5] In late afternoon, Captain Thomas Weir moved Company D from its position to what is now known as Weir Ridge about a mile away. This was contrary to his orders and occurred about the time the battle was ending. General Custer and 268 of his men were dead, and 55 wounded.[6] There was one Crow Indian scout working for Custer's command who escaped by feigning death. At least one, and probably two or more, of the Army personnel disobeyed orders or fled during the battle.

---

[3] "Battle of the Alamo," Wikipedia Encyclopedia.org, http://en.wikipedia.org/wiki/Battle_of_the_Alamo

[4] "Battle of the Alamo," Thomas Legion.net, http:thomaslegion.net/battle_of_the_alamo.html

[5] "The Battle of the Little Bighorn, 1876," Eye Witness to History, http://www.eyewitnesstohistory.com/Custer.htm

[6] "The Battle of the Little Bighorn," Wikipedia Encyclopedia.org, http://en.wikipedia.org/wiki/Battle_of_the_Little_Bighorn

Thus counting two fleeing out of 269 dead and 55 wounded, the incidence of trust violation in combat was 6×10^{-3}.

Rorke's Ridge – The senior British commander in southern Africa, Lieutenant-General Lord Chelmsford, invaded Zululand in 1879 to break up the Zulu political and military system. His strategy, framed by High Commissioner, Sir Henry Bartle Frere, was based on the conviction that the Zulu kingdom was a block to Confederation and wanted it removed to suppress African resistance. Seizing on a number of border violations and long-standing boundary dispute with the Transvaal, Frere presented the Zulu king, Cetshwayo kaMpande, with an ultimatum in December 1878. The ultimatum demanded that Cetshwayo disband his rebel force, cease requiring tribute from young men through military and social service, and to surrender authority to a British resident. It was a demand, as intended, no independent ruler would accept. On 11 January 1879, the ultimatum expired and Lord Chelmsford's troops invaded Zululand.[7]

Initially the operation went well. Chelmsford established a camp on the forward slope of a mountain, a location that commanded a view of several miles of open country. Although he felt secure about the left flank and front, he was concerned about hills on his right. Beyond the hills the country consisted of ridges and steep valleys where a Zulu army could move unopposed. On 21 January, he sent African auxiliaries and most of his mounted men to search the hills. They found a strong Zulu force and reported this to Chelmsford.

Chelmsford ordered half the command (the 2/24th with four of his six guns) to march out of camp immediately. His intention was to surprise the Zulus at dawn. He left the 1/24th to guard the camp, and ordered one of two support columns up Rorke's Drift. The camp was left under command of Lieutenant Colonel Henry Pulleine. The contingent, one company of 139 men, sent up to Rorke's Ridge is the focus here.

A Zulu force of 4,000 men attacked the line of British troops protecting the base camp. They broke up defensive formations, using spears and rushing into the camp. Similarly, on the Ridge above the camp, they attacked in overwhelming numbers. At the height of the battle, the killing was intense as soldiers, unable to escape, fought on with clubbed rifles, fists, knives, and even stones. "Those red soldiers," recalled one warrior, "how few they were, and how they fought; they fell like stones, each man in his place."

Lieutenant John R.M. Chard, in command of the company on the Ridge, left a brief report indicting that Lieutenant Adendroff, accompanied the troop deployment, was to stay. But, there is overwhelming evidence that Adendroff rode off shortly after their arrival. He was later arrested in the village of Pietermaritzburg and charged with desertion in the face of the enemy. He was the only one to escape. There is no evidence that a trial was ever held.[8] Of the total 1,700 men in

---

[7] "The Battle of Isandluwana," Ian Knight, Isibindi Africa Lodges, World Wide Web, 4 pages. http://www.zulunet.co.za/izl/isndlwana.htm.

[8] "Zulu War 1879 Discussin & Reference Forum (A Small Victorian War in 1879): Lt. Ardendorff-Deserter?, World Wide Web, 2 pages. http://1879zuluar.talk-forums.com/t2096-lt-ardendoff-deserter

the British camp at the beginning of the campaign, over 1,300 were killed. This yields an incidence rate of one in 1,700, or 6 x 10^{-4}.

These 19th century small action cases are a sample of many such throughout the past. Most are unrecorded, and those that are suffer from gaps and ambiguities in the accounts on both sides. But the trust violation rate seems to be roughly 10^{-3}.

TRUST VIOLATIONS IN THE FEDERAL JUDICIARY

Within this broader view of trust violations vice the earlier focus on espionage in Report 2, we return to the case of William M. Merrick, a Federal Circuit Court judge in the District of Columbia during the Civil War. He was appointed to the bench by President Franklin Pierce in 1855. From the beginning of the war, Merrick was suspected of disloyalty to the Union and favoring Washington taken over by the Confederacy. He refused to administer the oath of office to one of Lincoln's appointees to the Treasury Department, telling the man, "he was unfit to hold office." The Secretary of the Navy, Gideon Wells, noted, "the hearts and sympathies of the present judges in Merrick's court are with the rebels."[9]

Merrick's issued writs of *habeas corpus*, demanding underage boys who had enlisted in the Union Army be sent home. He issued about twenty of these writs in late summer and autumn of 1861 and granted orders, citations, and encouraged lawyers to harass defendants and delaying hearing the cases. On 21 October 1861, Secretary of State William H. Seward directed Brigadier General Andrew Porter, Provost Marshal in Washington, "to establish a strict military guard over Merrick's residence." Merrick was not considered under arrest. President Lincoln ordered that Merrick's pay be suspended, "to make him understand that at a juncture like this, when the enemy is at it were at the gates of the capital, the public safety is deemed to require that [Merrick's] correspondence and proceedings should be observed."

Merrick considered himself under house arrest, sent letters to two other judges to protest General Porter's actions as obstruction of the law. He asked Judge Morsell, "If martial law is to be our guide, we took the President of the United States to say so." He was told, "The privilege of the *Writ of Habeas Corpus* has been suspended for the present by order of the President." This case was without parallel in U.S. judicial history. Upon review, it was found that Lincoln never publicly or officially suspended the Writ of Habeas Corpus. Further, many believed that privilege was a responsibility of Congress and not a presidential power.

When the Union separated in 1861, most Federal District Judges living in Confederate states were reappointed by President Jefferson Davis to retain their positions. This included 13 Districts, each with about 13 judges, for a total of 169 judges, that joined the Confederacy. The Confederacy had no Supreme Court.[10] The nine Federal Supreme Court Justices remained with

---

[9] "Sweltering with Treason: The Civil War Trials of William Matthew Merrick," Jonathan W. White, 2007, World Wide Web, 9 pages. http://1879zuluar.talk-forums.com/t2096-lt-ardendoff-deserter

[10] "Confederate States of America," Wikipedia Encyclopedia.org, World Wide Web, 8 pages. http://en.wikipedia.org/wiki/The_Confederacy

the Union. The 23 states and District of Columbia, which remained with the Union, had 27 District Court Federal Justices, each also with about 13 judges, for a total of 351 judges, and four U.S. Court of Claims judges appointed by Lincoln during the war.[11] Thus, the primary Federal courts in the Union and Confederacy together had about 533 judges.

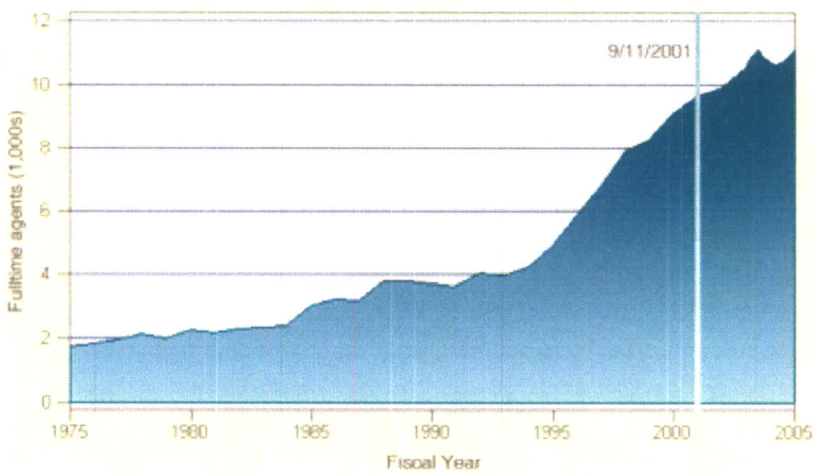
Based on this single case, the incidence of trust violations among the Judiciary uring the Civil War was 2×10^{-3}.

TRUST VIOLATIONS IN BORDER CONTROL AND CUSTOMS INSPECTION PERSONNEL

This class of trust violations include accepting bribes for permitting smuggling, allowing illegal immigrants entry, assisting in moving of drugs or other illegal contraband across the border, and stealing valuables from passengers or tourists coming into the country legally.

Figure 2 shows the full time Border Control and Customs field agents over time. Several sources are available. The second most important measure is the number of agents arrested and charged with corruption each year. The number of arrests for corruption for fiscal years 2004, 2005, and 2006 are from a Washington Post report.[12] Another report indicates that according to DHS figures, 129 officers were arrested on corruption charges between 2003 and 2009.[13] With these data, it is possible to construct Table 3 for fiscal years 2004-2009. A few numbers were computed by subtracting the known year arrests from the summation of 129 arrests between 2003 and 2009.

Figure 2
DHS agents deployed along borders and ports of entry



[11] "Judges Appointed by Abraham Lincoln," Wikipedia Encyclopedia.org, http://en.wikipedia.org/wiki/List_of_federal_judges_appointed_by_Abraham_Lincoln

[12] "Bribery At Border Worries Officials," John Pomfret, Washington Post, 15 July 2006, pg A1. http://www.securitycornermxico.com/index.php?option=com_content&viewarticle&id=1248&catid=60&itemid=1223

[13] Border Patrol Agent Arrested on Marijuana Charges," Reuters News, Phoenix, http://www.reuters.com/article/2011/04/05/us-crime-borderpatrol-idUSTRE73483B20110405

Table 3
Border force sizes and corruption arrest reports: 2004–2009

Fiscal Years	Full-time Field Agents	Agents Arrested	Incidence rate for arrest
2003	9,902	??	---
2004	11,106[14]	~ 26 (comp.)	2×10^{-3}
2005	11,200 (est.)	22[12]	2×10^{-3}
2006	13,000 (est.)	21[12]	2×10^{-3}
2007	15,000 (est.)	~ 20 (comp.)	1×10^{-3}
2008	17,399	17[12]	1×10^{-3}
2009	19,2301[5]	23 (comp.)	1×10^{-3}

TRUST VIOLATIONS IN TRANSPORTATION SECURITY AGENCY INSPECTION PERSONNEL

Thefts include cash, travelers checks, perfume, and jewelry. More recently, electronic devices such as laptop computers, computer games, GPS units, and cell phones have been added to the list. In August 2004, there had been 28,000 claims of loss or damage filed to the agency for the first two years of operation. In 2009, they were down to 11,700 claims.[14]

The following data is to summarize the published reports:

- In January 2002, TSA employed 1,300 FAA baggage screeners and began the new training program. By May, those personnel were ready for deployment as TSA hired thousands more federal employees, with a goal of obtaining 28,000 screeners by the beginning of 2003.[15] During 2003, TSA fired 1,200 employees after they failed basic criminal background checks.[16]
- By November 2004, TSA had a force of 45,000 but downsized in two steps: 3,000 by 31 May 2005 and 3,000 more by 30 September 2005. Despite a Congressional cap of 45,000 agents and the reductions in force, TSA hired 10,000 "temporary help," for a net increase of 4,000 people to 49,000 by the end of 2005.[17]
- In December 2008, some TSA baggage screeners were rotated back to upgrade security training through 31 July 2009. At the time the force was estimated at 50,000 personnel.[18]

---

[14] "Theft and Corruption: Just Another Day at the TSA," Foolish Nation, http://www.foolishnation.com/
[15] "Aviation Security Agency Gets Off the Ground," Government Executive 21 December 2001, http://www.govexec.com/dailyfed/1201/122101p1.htm
[16] "Theft and Corruption: Just Another Day at the TSA," Foolish Nation, http://www.foolishnation.com/
[17] "Airports Focus On Funding For Baggage Screening Equipment," Archive, 9 May 2003, http://archives.californiaaviation.org/airport/msg26140.html
[18] "TSA Screening is Security Theater," Lesley Stahl, CBS News "60 Minutes," TV interview.

- In June 2011, the agency employed about 60,000 screeners, including an additional 10,000 baggage and passenger screeners holding security clearances.[19] In June 2011, the agency employed about 60,000 screeners (baggage and passenger screening).[20]

These data are shown in Table 4:

Table 4
TSA force sizes and corruption arrest reports: 2002–2011

Calendar Years	TSA Force Size	Agents Arrested, or Fired	Incidence rate arrests or fired
Jan.- May 2002 31 Dec. 2002	1,300 < 27,800	0 ~ *25*	 9×10^{-4}
May 2003	28,000	~ *35*	1×10^{-3}
Nov. 2004	45,000	60	1×10^{-3}
31 May 2005 31 Dec. 2005	~ *48,000* ~ *49,000*	------- ~ *70*	 1×10^{-3}
Jan. 2006	50,000	-------	------
2007	< *49,000*	~ *70*	1×10^{-3}
2008	50,000	~ *73*	1×10^{-3}
2009	50,000	~ *75*	1×10^{-3}
Sept 2010	< *60,000*	~ *75*	1×10^{-3}
June 2011	~ 60,000	> 48 to ~ 65	1×10^{-3}

The estimates in the third column (noted by italics) were generated by distributing the TSA 2009 total estimate in 2009 of 500 arrests plus 48 verified cases in Honolulu in 2011 (total of 548.)

There is a wide variation in the cases above. Stealing a bottle of perfume is not as serious as more recent substantive thefts.[21]

- Pythias Brown: Stole electronics worth tens of thousands of dollars (Oct. 2008)
- Al Raimi: Stole $30,000 from travelers and bribed his supervisor to ignore it (Oct. 2010)
- Michael Arato: Took bribes from fellow workers who stole $30,000 (Oct. 2010)
- Davon Webb: Stole $200,000 from passengers (Feb. 2011)
- Nelson Santiago: Stole $50,000 worth of electronics from a passenger's baggage (July 2011)
- Karla Morgan: Arrested for theft of $1,000 in "lost wallet" sting (June 2011

[19] "TSA Theft of Passenger Valuables a Nationwide Problem," Howard Portnoy, Hot Air, http://hotair.com/archives/2011/06/20/tsa-theft-of-passenger-valuables-a-nationwide-problem/
[20] "TSA Theft of Passenger Valuables a Nationwide Problem," Howard Portnoy, Hot Air, http://hotair.com/archives/2011/06/20/tsa-theft-of-passenger-valuables-a-nationwide-problem/
[21] "Theft and Corruption: Just Another Day at the TSA," Foolish Nation, http://www.foolishnation.com/2011/07/08/theft-and-corruption-just-another-day-at-the-tsa/

TRUST VIOLATIONS IN TERRORISM AND VIOLENT CRIME WORLDWIDE

It is instructive to look at statistics relating to the fraction of people in a group who resort to violent behavior. Two such examples are people who become terrorists and the those who commit violent crimes. The following Table 5 suggests that the percent of people electing illegal violence is of the order of 10^{-3}. The terrorist countries are a representative set currently experiencing terrorist activity. The terrorism statistics are from the IISS tabulation of sub-state armed groups.[22] The violent crime statistics are for 2000, drawn from an EU report.[23] Yellow shading indicates those countries where there are data for both terrorism and violent crime.

Table 5
Terrorism and violent crime rates in various countries

COUNTRY	INCIDENCE OF TERRORISM	COUNTRY	INCIDENCE OF VIOLENT CRIME
Ivory Coast	1×10^{-2}	Bulgaria	8×10^{-3}
Liberia	7 xo 10^{-3}	Norway	4×10^{-3}
Iraq (AQ only)	4×10^{-3}	Hungary	3×10^{-3}
Chechnya	3×10^{-3}	Russia	6×10^{-4}
Corsica (France)	2×10^{-3}	France	4×10^{-3}
Palestine	2×10^{-4}	Denmark	3×10^{-3}
U.K./Muslim	1×10^{-3}	U.K.	1×10^{-2}
U.K./Northern Ireland	8×10^{-4}	U.S.	5×10^{-3}
Columbia	6×10^{-4}	Canada	9×10^{-3}
Afghanistan	6×10^{-4}	Germany	2×10^{-3}
Philippines	4×10^{-4}	Poland	2×10^{-3}
Turkey	1×10^{-4}	Turkey	1×10^{-3}
Algeria	1×10^{-4}	Portugal	2×10^{-3}
Pakistan	1×10^{-4}	Italy	2×10^{-3}
Japan	4×10^{-5}	Japan	6×10^{-4}
Peru	2×10^{-5}	Greece	8×10^{-4}
Australia	1×10^{-6}	Australia	9×10^{-3}

The two sets of data are not directly comparable but they serve to give an idea of the proclivity of people to move outside legal norms and to participate in violence. The terrorism numbers are admittedly uncertain estimates of group size, and not all these individuals will have personally committed, though they will have supported, violent acts. The violent crime numbers vary with individual national definitions of the term, and suffer from variations in reporting rates. In the

[22] International Institute for Strategic Studies, *The Military Balance: 2006,* London
[23] See www.csdp.org/research/hosb1203.pdf

U.S., violent crime is defined as murder and non-negligent manslaughter, forcible rape, robbery, and aggravated assault. Furthermore, since this is a count of acts, not people, the numbers do not reflect the number of acts per individual and the number of individuals per act. The predominance of terrorism and crime varies. The terrorism participation rate is higher in Chechnya than it is for crime in Russia overall; in the U.K., Turkey, Japan, and Australia the crime rate is higher than the terrorism rate by an order of magnitude.

TRUST VIOLATIONS IN TERRORISM IN THE U.S.

A recent report described the change in the FBI's investigation process following President George Bush's decision to relax limits on domestic-intelligence aimed at finding terrorists in the U.S.[24] The data began in December 2008 and extended through March 2009. During that period, the FBI made a several-times increase in number of investigations that it initiated, to a total of11,667 total. Of these, 8,605 investigations were completed and apparently found no terrorists based on low-level inquires. An additional 427 cases were set aside for intensive investigations. The disposition of the remaining 2,635 cases was not indicated and there were no data from the years before the rules changed to compare against. Agencies typically do not provide statistics interviews or investigative cases terminated for various reasons.

The fact of 8,605 cases found no evidence of terrorism suggests that the terrorism incidence rate in the U.S. population is less than 1×10^{-5}.

CONCLUDING OBSERVATIONS

The final data are shown in the discussion to one significant figure only to express the view that they are not worth greater precision. Statistics are wonderful but, in a quotation of Josiah Stamp, statistician and Director of Inland Revenue in the U.K.:

"The government are very keen on amassing statistics. They collect them, add them, raise them to the nth power, take the cube root and prepare wonderful diagrams. But you must never forget that every one of these figures comes in the first instance from the village watchman, who just puts down what he damn pleases."

Matters of definition, local culture, quality of statistical agencies, and the degree of record-keeping at various times and circumstances affect the numbers. There is no significance to be attributed to the difference between 1 and 1.4 when the rates are distributed over four orders of magnitude. Further, focusing on the exponent, essentially the log of the incidence rates, makes it simpler to see what the data are telling us.

By presenting incidence rates. i.e. a fraction of the people in a population to which the identified person is seen to "belong" implies we believe the number of people displaying a common characteristic such as espionage, terrorism, violent crime, theft, etc. varies directly with the size

[24] "FBI Casts Wide Net Under Relaxed Terror Probe Rules," Charlie Savage, New York Times News Service, The San Diego Union Tribune, 27 March 2011, page A7.

of the parent population. There are three hypotheses hidden in this. First, the incidence of the selected behavior is relatively infrequent. In a society lacking norms of behavior, incidence rates are meaningless. The behavior selected has to be sufficiently unusual that we choose to characterize it as a behavioral anomaly worthy of data collection and analysis. The second hypothesis is that we believe all incidents of a "type" are the same: X's act of espionage is the same as Y's act of espionage. But we know nothing about X and Y. Even the most detailed interrogation and investigation by law enforcement, intelligence, and psychological professionals may not know if what X did is the same as Y. The similarity lies in the arbitrary definition of the act as defined by legal or cultural norms. And third, presenting a "rate" implies the we have collected all, or a substantial fraction of the occurrences of the behavioral characteristic we identify, and that we know the size and composition of the parent population from which these individuals are presumed to be a member.

To go forward, the reader must at least stipulate that he *tentatively* accepts the above three hypotheses as valid. This said, what can we say further? In the 1 October report we outlined how RTISFS could be operated to look for evidence of deception by analyzing the results of several cycles of textual query–responses initiated by a user activating a software sensor set to trigger on an action seen as possibly part of a violation of trust to protect sensitive information.

Recall why the incidence rate of what we are looking for is central to indicating its presence. When the system is installed, it is a blank slate. There are N users about whom we know nothing. Yes we have their life record, the things they have done since their employment at the facility installing RTISFS, and such intuitive judgments as security personnel may already have. For our purpose we treat these as prejudices based on matters that may not even be correct.

Where, then, do we start. If we can avail ourselves of relying on an incidence rate of 10^{-3}, which is indicated by the data above, we can reliably use samples of 1,000 users to establish baselines from which possibly anomalous behavior can be judged, secure in the knowledge that 99.9% of our "training sample" is not deceptive. As the expected incidence rate increases above 10^{-3}, we can only reliably use smaller samples to set baselines and with such small samples, our baselines will be increasingly unreliable as indicators or anomalous behavior.

So what does our admittedly "quick-and-dirty" review of the trust violation literature tell us? Consider the seven sets of behaviors that have been examined: Espionage since 1770; desertion in military conflict; disloyalty by the Federal judiciary in the Civil War; corruption in field unity of the Department of Homeland Security. Current terrorism, both worldwide and in the U.S.; and violent crime worldwide and in the U.S.

We come to the following conclusions:

Any group is reliably trustworthy, to 99–99.9%;

> In cases where the penalty for untrustworthiness, or the likelihood of being caught, is seen by the offender as tolerable or not serious, the likelihood of trust violation in the group is about 10^{-3}.

In cases where the penalty for untrustworthiness is seen by the offender as serious, serious enough to take efforts to prevent discovery, the likelihood of such a violation is about 10^{-4}.

Psychopaths are one group that may not fit these conclusions. There may be other classes of people whose world view is such they do not fit into these incidence expectations either.

The data on which these conclusions are based are are presented in Table 6:

Table 6
Trust violation incidence rates for six major categories examined

MAJOR CATEGORY	SUBCATEGORY	DECADE(S)	NO. OF CASES	INCIDENCE RATE	PERCEIVED SERIOUSNESS
Espionage	U.S. Revolution	1770	10	7×10^{-4}	Incr. not
	U.S. Civil War: Union	1860	3	1×10^{-4}	
	U.S. Civil War: Confed.	1860	3	5×10^{-7}	
	WW II Venona +	1940	600	2×10^{-4}	
	Cold War	1950-2000	44	1×10^{-5}	
Military combat	Alamo: Texas	1830	1	4×10^{-3}	
	Alamo: Mexico	1830	1	4×10^{-4}	
	Custer's Last Stand	1870	2	6×10^{-3}	Not
	Rorke' Ridge (Zulu War)	1870	1	6×10^{-4}	
Federal Judiciary	U.S. Civil War	1860	1	2×10^{-3}	Not
Homeland Security	Border Control	2000	129	$1–2 \times 10^{-3}$	Not
	Transportation Security	2000	488	1×10^{-3}	Not
Terrorism	Worldwide	2000	>10,000	$10^{-5}–10^{-2}$	
	U.S.	2000	8,605	$<10^{-5}$	
Violent Crime	Worldwide (incl. U.S.)	2000	Unk.	$10^{-3}–10^{-2}$	Not

RUSSIAN FOREIGN INTELLIGENCE

Fritz W. Ermarth

Although the Cold War ended more than twenty years ago, Russian intelligence poses a challenge to U.S. intelligence, counterintelligence, and national security of considerable importance and severity. This has been the case for most of the past twenty years and appears likely to continue for the indefinite future. The gravity and scope of this challenge is outranked only by that of China.

LEGACY FACTORS AND ELEMENTS OF CHANGE

Russia, like the USSR, runs a broad spectrum intelligence effort against the United States. It aims to collect both protected (classified) and unprotected information about U.S. politics, economics, business, technology, military strength and programs, U.S. intelligence, and personalities. It applies all the traditional disciplines of human and technical intelligence, with a heavy traditional emphasis on humint (recruitment and operation of espionage and influence agents.) But it adds a heavy emphasis on a new discipline, a powerful evolution of the old discipline of sigint: intelligence and influence action operations to have an influence in cyberspace.

The legacy of a centuries-old political culture is important for Americans to grasp. Russia is a low-trust society in which coercion, manipulation, and intrigue have long been regarded as more important means to accomplishment of goals among people, communities, states than overt bargaining, negotiation, and adjudication. This attribute has given clandestine intelligence activity an importance in Russian statecraft probably greater than that of any other major power in modern history. This was a cultural quality that made the imposition of the combative and conspiratorial ideology of Marxism-Leninism-Stalinism rather easy and natural. Post-communist Russia has, in many ways, reverted to pre-communist cultural norms that have reinforced this cultural legacy.

Like the USSR, post-communist Russia harbors a multiplicity of agencies with intelligence and counter-intelligence missions and activities. Legatees of the old KGB: the SVR for foreign intelligence and the FSB for domestic security and counterintelligence, are institutionally and politically the most important. The intelligence arm of the Russian General Staff, the GRU, although currently under a kind of attack for domestic political reasons, remains large and important. And there are a variety of other agencies for intelligence and counter-intelligence in the areas of narcotics, terrorism, electronic security, customs, taxation, etc. All this might be likened to the large and diverse intelligence community of the United States. But in the case of Russia, there is less community and far less subordination to the rule of law.

Like its historical predecessors, contemporary Russian intelligence ascribes a great important to influence operations and activities, that is, the development and use of agents and channels that

can be used, not merely for informational advantage, but to directly influence the behavior of targeted states or other entities.

In Soviet times, going back to the revolution of 1917, a major channel and arena of influence, as well as intelligence gathering, was the international communist movement. This has largely dissipated. But not entirely. Agents and channels derived from that movement are still in existence in most parts of the world for use in serving contemporary purposes, political, strategic, and commercial.

Post-communist Russia has discarded the universalist ideology of Marxism-Leninism which aimed to extend communist rule throughout the world. But those who preside over Russia's foreign and national security policies have retained, or reverted to, the age-old mission of making Russia a great power in the world and of restoring some semblance of the old Russian empire on the territory of the former Soviet Union and Soviet bloc. And this is not entirely different from how Soviet leaders came very early to identify the augmentation of the power of the USSR as the most important aspect of the international communist mission.

During the Cold War, the specter of a possible nuclear war between the U.S. and the USSR was a central focus for Soviet intelligence, influence operations, and counter-intelligence. Gathering intelligence that would help keep this specter at bay or give the USSR advantage should it arise was a top priority. This priority has subsided, but not disappeared by any means. Russia assigns a very central role to nuclear weapons in its national strategy and is trying to modernize its forces to underwrite this role, by providing both a central, survivable strategic nuclear deterrent and an array of more limited, agile, usable nuclear forces for tactical or battlefield and discriminating strategic use. To assure the viability and credibility of options in the latter class of capabilities is the major reason for tenacious Russian opposition to U.S. plans for limited ballistic missile defenses, especially in Europe. For the same reasons, the Russians are determined to do what they can to dissuade the U.S. from modernizing its own nuclear arsenal and, thus, are exceptionally secretive about their own modernization and energetic in lobbying against comparable U.S. efforts. This preoccupation keeps matters pertaining to nuclear weapons strategy and policy at a high priority for Russian intelligence. And it very much influences Russia's approach to arms control issues.

In this context, it should be noted that Russia shares U.S. appreciation of the importance of nuclear proliferation as a challenge to national security. But its attitude is more nuanced and differentiated because it sees, rather as China does, nuclear proliferation as less a threat to itself than a challenge to the primacy and dominance of the U.S. in world affairs, an attitude that very much influences its policy on Iran.

The USSR, like the U.S., ran a world-wide intelligence effort because the central deterrent standoff made far-flung regions into arenas for the conduct of that conflict, in addition to their intrinsic importance.

The end of the Cold War greatly reduced the importance of the Russian target for U.S. intelligence and resulted, in the eyes of some critics, in an excessive reduction of attention to that target. And, for the U.S., there very quickly arose a different landscape for apportioning

intelligence resources and attention, a landscape of many, different, more equally competing intelligence priorities in proliferation, terrorism, crime and narcotics, money laundering, the geopolitics of energy, etc.

Russian intelligence faced a similar challenge. Old priorities retained their importance. But others gained new and competing importance. The new issue profile facing the U.S. also faces Russia. But of special, unique importance to Russia is its so-called Near Abroad, the newly independent states of the former USSR, territories of the former Russian empire. Gathering intelligence and exerting influence on these countries has a priority for Russia that, by comparison, the nearby regions of Latin America have never attained for the U.S. Old ties of party, KGB, military, and business industry origin provide Russia with strong platforms for the conduct of intelligence and influence operations. The affect of these ties has been vividly and faintly on display all around Russia, in Ukraine, Georgia, the Baltic countries and elsewhere At issue are the contest for influence with competitors, especially China and the U.S., the detection and aversion of new threats such as Islamic fundamentalism and terrorism, the struggle over control of energy resources, and the prospects for reconstructing something like the old Russian empire in the guise of a new Eurasian community or the like. Notwithstanding the resource and budget windfall from energy export that has greatly helped to revive Russian intelligence since the collapse of the USSR, these interests are a weighty tax on the whole Russian intelligence effort.

A new factor in Russian intelligence, of great importance in scale, priority, and diversity is simply that of doing business and making money. This point was made at a recent meeting of international political and business figures, by a former officer of the GRU, now in private business intelligence work. Asked, "If you were in charge of all Russian intelligence, what would be your priorities for understanding the U.S.? The Russian responded, "Always keep in mind that the top priority for all Russian intelligence agencies and people and their masters is to make money." It became clear that this interest or priority, while often overlapping with, is distinct from traditional interests of state. It is animated by private, personal, group, subgroup, and clan interests. Although novel in scale, priority, and distinction from state interests as usually conceived, this interest of Russian intelligence in business and money has a long and relevant history.

Recall the USSR's creation of such organizations as Amtorg in the 1920s to serve as platforms for the deployment of non-official cover agents, illegals, and the collection of commercial and business intelligence as well. In the mid-1980s, Soviet intelligence was directed by high political authority to expand and diversify the creation of commercial platforms for intelligence operations against the usual military and political targets, but also for the collection of technology, commercial, and financial information. As Gorbachev opened avenues for private activities, many Russian intelligence entities were positioned to take advantage of this for private or parochial gain abroad. There began a tsunami of making and moving money. One major objective was the creation of overseas financial and physical refuges for members of the Soviet elite who feared dislodgment by the impending revolution. A major result was the deep and growing entrenchment of Russian intelligence in overseas business, very often shady by Western legal standards, quite often in league with Russian and other organized crime entities. Energy resources, other raw commodities, and arms were the primary wealth generators..

This phenomenon flowered greatly with the collapse of the USSR and communist rule in Russia. It has become an important aspect of Russian domestic and international affairs. A significant, albeit temporary, irony of this"takeoff" period of the privatization of the post-communist Russian state was that the very centrally positioned elements and elites of Russian intelligence, especially the KGB, were not the primary beneficiaries initially. The advantage went to a small class of "appointed capitalists" drawn from the periphery of the Soviet party elite (nomenklatura) who quickly assembled business empires of great wealth and became the politically powerful and deeply resented oligarchs of the Yeltsin period. A feature of this class, especially resented by those who benefited less than they thought they deserved, was that many of them were Jews.

The appointment by Yeltsin of Vladimir Putin, a little known functionary in his entourage, as successor to the Russian presidency in late 1999 and the latter's consolidation of personal power occasioned a major "correction" to this state of affairs. Putin, himself a veteran of the KGB's First Chief Directorate for foreign intelligence, at least knowledgeable about money operations from his perch in East Germany in the 1980s, and a player in such activities during a position under the mayor of St. Petersburg later, put KGB officials and veterans into powerful positions in the echelons of the state apparatus and somewhat renationalized state-scale enterprises, especially in the energy sector. Some oligarchs went into exile (most prominently Boris Berezovskiy). One, Mikhail Khodorkovskiy, went to jail and his business empire was apportioned to new oligarchs mainly because this figure sought to use his wealth for politics, aiming especially to bring Russia under the rule of law and Western business, legal, and political norms. The result of Putin's policies and personnel moves was vastly greater power and legitimacy for Russian intelligence entities to conduct business on the international scene. This new condition was welcomed by its beneficiaries. But it is not untroubled.

THE "STATE" OF RUSSIA

Russian intelligence is not insulated from conditions and politics inside Russia. In Soviet times, Russian intelligence agencies, especially the intelligence arm of the NKVD, were deeply involved in Stalin's struggle to eliminate Trotsky and all traces of Trotskyism inside and outside of the USSR. All became victims to some extent of the Great Purge.

Today, the condition, resources, priorities, behaviors of Russian intelligence entities are very much influenced by the "state" of Russia, that is, the condition of Russia as a state, country, and society. This deserves some examination, both to shed light on current behaviors as well as to anticipate future developments which could be dramatic.

According to the characterization of a U.S. diplomat serving in Moscow, as revealed by a recent account derived from Wikileaks, Russia is a Mafia state. This is a thumbnail depiction of a complex phenomenon. The label is meant to convey that Russia is ruled by an elite that is made up of clans of power, influence, and money that are both the beneficiaries of state power and the foundations of it. They and the individuals who run them compete, sometimes vigorously, but are united in dependence for wealth and even survival on a system generally called the Putin regime.

Both formal and informal authority relations, including very important personal and blood-family relations, structure this system and shape its dynamics.

The mafia metaphor denotes power and wealth arrangements observed in many countries , particularly in the Third World. It also carries a connotation of criminal features quite apt in the case of Russia. The business and power structures of Russia involve some participation by and cooperation with organized crime, and very definitely the engagement in "authorized crime" by entities in that structure particularly bribery, extortion, embezzlement, and often violent attacks on perceived adversaries. These adversaries may be participants in the system such as rival clans and business entities. Or adversaries of the system, dissidents such as journalists and outside-the-system politicians who are battling it by exposes and arguments on behalf of another system, a liberal, law-governed, truly democratic system or, in a few cases, a return to outright Stalinism.

Both of these "faces" of mafia Russia, clan politics and criminality face outward as well as inward. As jailed oligarch Khodorkovskiy has observed, Russia has two exports: energy and corruption.

This mafia system needs a "don of dons" to preside over it. And it has one in Vladimir Putin, the past and future president of Russia. He inherited the elements of this system from Yeltsin. He consolidated them and restaffed its leadership, establishing rules of behavior, mostly that wealth and survival require acceptance of the system that made for the quasi-stability attributed to Russia under his regime. He and his inner circle have great but not unlimited power. Put in other terms, his is an authoritarian system with less authority than meets the eye. The essence of this system is the nexus or interdependency of power (political, institutional, business) and wealth. The main power of the system is its ability, so far, to quash open political challenge and to apportion wealth within the ruling elite.

In terms of political science and history, a more elegant label for Russia's system might be Byzantine Financial Feudalism. We all understand what feudalism is from the example of medieval Europe: An authority structure where vertical and horizontal power relations are defined by both formal/legal, symbolic/traditional, and informal "personal arrangements" with a connotation of force as required. In the Middle Ages, the economic base for feudalism was land and peasants. In contemporary Russia it is control of money flows based on energy, other commodities, arms, and real estate; hence financial feudalism. It is Byzantine, that is Eastern, in that it is not governed, as was Western feudalism, by the rule of law. The rule of law tradition, arising in ancient times from the Old Testament, Athenean Greece, and Republican Rome was transmitted into Western Europe of the Middle Ages by the Roman Catholic Church. This was not done by the Eastern or Byzantine Church. So the rule-of-law tradition is weak to absent in the political cultures of Russia and southeastern Europe. And the feudalistic system now regnant in Russia is perhaps most characterized by that weakness. But this also says that the way Russia is ruled today is rooted in ancient cultural traditions that will not be easily supplanted. And it helps explain how, after the collapse of the USSR, Russia developed under Yeltsin and Putin into a mafia state, rather than develop an authentic law-governed democracy.

Nevertheless, the state of the world, especially its information and knowledge conditions, demands that the Russian system pay some homage to the values of democracy and rule of law.

Hence, in ways not authentic, but more authentic than in Soviet times, the drama of election and voting must be conducted, and carefully controlled, on behalf of the system. A round of such dramas is impending in the December 2011 elections to the Duma and the March 2012 election of the president. Hypocritical observance of the rule of law is evidenced in the elaborate court proceedings that put and keep Khodorkovskiy in prison and many other cases.

There is mounting evidence of late that the long-patient Russian people, leave aside intellectual and political elites of liberal democratic persuasion, are becoming dissatisfied with this regime and its pretenses, notwithstanding the superficial stability and modest increase in general welfare it has brought. This is because of the growing urgency of a deep problem ever recurring in Russian history: the need for and deep difficulty of economic modernization.

To make a large and long story short, Russia's economic condition is defined by dependence on the revenues of oil and gas, and an archaic and decaying industrial base left by the collapsed USSR, which was archaic and decaying well before the USSR collapsed. Everyone from Putin to the man in the street knows that to survive as a country, not to mention as a power in the world, Russia must execute a great leap forward into a world where much of the nation's wealth and growth arise from the creation and use of modern technology, from socio-economic innovation, from domestic entrepreneurship, from efficiency in the use of resources, from transition out of a culture of wealth extraction and diversion into a culture of wealth creation.

Medvedev, Putin's appointee as temporary president, made this modernization his public agenda. But it really has not gone anywhere because the system does not allow it to be done at all, especially under the pretended leadership of a largely puppet figure. The features of stagnation have continued: infrastructure decay, the demographics of declining Russian population from low birthrates to talent emigration, mounting corruption, and fiscal crisis hidden, not prevented, by energy revenue

The vast majority of the vocal challengers of the Putin system are liberal (in the old sense) democrats (also in the old sense). They insist that Russia's economic and social modernization requires a breakout in the direction of true democracy, that is, authentic political competition for accountable political power, and the true rule of law. This will allow conditions for modernization driven mainly by bottom-up innovation, investment, entrepreneurship. It will attract foreign investment too. Modernizers associated with the regime grope around for a state-driven model, more like that of China. But they confront the problem that all state-driven investment projects in Russian conditions are little more than money flows to be stolen by elements of the elite.

These conditions and the surrounding political frustrations are creating something of a crisis or pre-crisis for the Russian political system. The similarity to the period of stagnation, as last seen in the late Brezhnev years of the USSR, is widely remarked upon.

Many see the possibility of dramatic developments. Even the controlled and theatrically staged elections of the coming period could witness manifestations of the larger social frustration. As in the recent past, missteps by the regime on sensitive issues, like pension changes, could trigger

public outbursts. Small but vivid events like plane crashes or forest fires could trigger them. Or elements of the elite could lapse into open conflict.

A wide range of scenarios out of these conditions is deemed possible by Russian and foreign experts: A long continuation of the stagnant condition buoyed from time to time by high oil prices. The emergence of a Gorbachev to open the way toward modernization that is both top-down and bottom-up. Or the imposition of a truly authoritarian regime: Stalin with a modern face. Or the most feared scenario: the breakup of Russia into warring satrapies.

WHAT DOES ALL THIS MEAN FOR RUSSIAN INTELLIGENCE TODAY AND TOMORROW?

The main effect is to license and encourage activities that make money for individuals, groups, and organizations, but also on behalf of the state. As a result, intelligence and influence assets that help Gazprom and such entities are immediately more important than assets which might serve the Ministry of Defense. In the U.S., Wall Street is as important an intelligence target as our defense/military arena.

Another effect is to inject serious insecurity into the calculations of individuals and organizations. They have to be on the lookout for avenues of personal and financial refuge in case of major instability at home, much as in the late Soviet period.

But these conditions cannot be expected to diminish activities on the part of Russian intelligence that U.S. intelligence and counter-intelligence should regard as threatening enough to track and thwart. In the worst of times, Russian intelligence has shown a record of energetic activity. It is a matter of priorities and resources. And also that activities of perhaps secondary importance to Russian intelligence organizations can have serious consequences for the U.S. if they are conducted with any skill and perseverance.

All this means that U.S. intelligence and counter-intelligence should regard the Russian intelligence challenge as a combination, more complex and diverse than in the Soviet period, of threats or potential threats involving long traditional targets of intelligence and influence operations, e.g., high politics and diplomacy, military and related technology information, along with new threats in the domains of cyberspace and business.

U.S. VULNERABILITIES IN RUSSIAN EYES

As an open society, the U.S. has long been something of an easy target for foreign intelligence, Russian and others. But this same quality has also diminished the relative value of secret information, albeit making it especially valuable in special situations, e.g., early atomic weapons secrets or the codes divulged by Walker-Whitworth

The Soviet Union was, by contrast, a closed society and a very hard target. Good spies were hard to get and hard to run for the U.S. and its allies. But we recruited some and they were extremely

valuable, e.g., Pentkovskiy, Tolkachev, Kuklinsky. By the numbers, the USSR and its allies won the spy war of 1945–90. In a net assessment of value, however, the U.S. won it.

HOW MIGHT A NET ASSESSMENT OF INTELLIGENCE IN THE POST-COLD WAR TURN OUT?

One has to factor in that Russia regards the U.S. as a much more important national security and intelligence target/problem than does the U.S. regard Russia. We might be wise to recalibrate that judgment. An important aspect of this is the vulnerability or accessibility of the U.S. as an intelligence target in the post-Cold War era, as perceived and probably exploited by Russian intelligence, among others. The U.S. is, on balance, a more open and accessible intelligence target than ever, especially as one factors is in the huge window of cyberspace.

But there are social-psychological aspects as well. The end of the Cold War and the collapse of communist rule in Russia largely, albeit not entirely, removed the ideological garden in which the USSR found and cultivated numerous spies. But a somewhat compensating development has taken place in the U.S.. The removal of the threat of "global communism" and the prospect of a great nuclear war with the USSR has reduced the moral opprobrium that was associated with spying in general and for the Russians. Add to this the decline of adherence to traditional values of trustworthiness and honesty in many dimensions of life that grew out of the "cultural revolution" of the last decades of the 20th Century. All this has tended to reduce spying for a not-obviously-very-threatening foreign state from the level of treason, a great moral evil, to something more like insider trading, illegal, perhaps wrong in some lights, but not evil, unless you take seriously the larger consequences of widespread indulgence in this practice, which contemporary "how do you feel" ethics tend not to do. There are generational aspects to this, no doubt. But younger generations are, one might suspect, more vulnerable than older ones.

Russian intelligence is certainly mindful of these social-psychological-generational tendencies, especially in the human intelligence targeting. strategies it considers for delivering intelligence and influence value over the longer term.

THE RUSSIAN ILLEGALS PROGRAM AS A LONG-TERM INVESTMENT

These reflections ought to be kept in mind when pondering what Russian intelligence and its godfather Putin had in mind when, according to a number of sources, they ramped up their illegals program in the U.S. starting around 2000. This meant the systematic insertion of recruited, trained, and credentialed (that is, officially enrolled) Russian intelligence officers into the US with no official cover, but rather the cover of ordinary citizens, immigrants, and tourists for the advancement of intelligence objectives in most cases, the spotting and recruitment of American assets for those purposes, to be served from their present positions or in positions they could be expected to enter as their careers advanced.
This would fit naturally with the mentality and outlook of the people in charge of Russian national security during the past decade or so into the present. The U.S. is the main danger and will remain so. The U.S. is broadly accessible. A long-term investment of this kind is well worth it. Not very expensive: Some modest retainers to gain loyalty or fear of exposure. And not very

risky, especially if the Russian illegals so deployed do not themselves commit espionage and if their recruits are "bought" for deeds to be committed in the future.

The Russian illegals, "rolled up" in late June 2010, could have been a special part of this program, but a deceptive part, a venture in "maskirovka" or camouflage. They were decoys to distract and monopolize the attention of U.S. CI resources devoted to Russia, especially by the FBI.

Whatever its problems, the SVR is not operationally incompetent. The "sloppiness" of the illegals which aided in their surveillance was purposeful. Any trained operative would have early detected the heavy physical and technical surveillance they were under. Yet they continued their game for nearly a decade. Anna Chapman has to have known that the FBI agent to whom she handed her laptop for repair was not an SVR officer because she was under the control of a real one who would have told her. The handoff was intended to put a steganographic code into the hands of the FBI so it could read messages intended for them.

The eleventh illegal, one Metsos, the handler of the rest, escaped. The illegals committed no seriously prosecutable acts. This was under instruction. But they had to behave so as to keep attention on them, which they did.

All this leaves in question the actions, motives, and loyalties of the party who is supposed to have betrayed these eleven illegals nearly a decade ago and whose "defection" in late June 2010 triggered their arrests, as well as his Moscow "trial" in absentia. As is the question of what could have been learned from the arrestees had they not been returned to Russia so quickly.

There are two conclusions to be reached on the basis of this episode that are mutually incompatible. If the U.S. official version of this episode is valid, the SVR has become a supremely incompetent institution. In this case, one may cease worrying about Russian intelligence, or at least about SVR-run humint operations until further notice. Or the exposed and arrested illegals, probably deployed as decoys, were the fringe of a larger and much more serious operation, a long-term investment in the recruitment of potential future assets, exactly of the sort that call for detection on the basis of their own information gathering activity when they are activated.

A postscript on these speculations is in order, another vulnerability as perceived by Russian intelligence. Despite the decline of adherence to, and celebration of, traditional values, the U.S. remains a high-trust society. As individuals and groups, this tends to make us believe uncritically in what we observe or think we observe. That makes us highly vulnerable to sophisticated deception. If that deception is accidentally accompanied by observables that betray it, we are sharp watchers and may well pick them up. But we are quite disinclined to ask *ab initio*, how might the observables I see be intentionally false and misleading? And what are the implications if they are?

Thus, if, as is likely, the Russians have deployed a broad humint operation to produce a fairly large number of agent assets for future activation in exploitation of computer accessible data,

among other objectives, it is certain that they will use that agent network to insert deception and disinformation into our system.

RUSSIAN VULNERABILITIES

One of the implications of the condition of Russia depicted here is that Russia and its security organs, including intelligence, are themselves vulnerable. Russians need, and are looking for, ways to escape from or provide themselves future security in a very insecure environment. If the Russian intelligence challenge is perceived as a threat to us, penetrating their agencies and operations is a vital defense. The operational opportunities for doing so are inviting. The principal downside of escalating humint operations against Russia is largely political and psychological. When Russians spy on us, we tend to write it off as Russians being Russian. When we spy on Russia, they tend to see it as another sign of deeply embedded U.S. and Western hostility to Russia, stimulating the nationalist paranoia that is part of the larger pathology of Russian political culture we, and Russia's own real democrats, hope to overcome. So even if we give a higher priority to the Russian intelligence and counter-intelligence challenge, exploiting Russian vulnerabilities requires great discretion and selectivity.

DO NOT FORGET INFLUENCE OPERATIONS

The primary purpose of the project for which this essay is written is to develop concepts and paradigms for the development of tools to detect espionage conducted on computer networks. But we must not forget that the larger Russian intelligence challenge includes the very important domain of influence, that is operations and relationships designed to encourage behaviors in the U.S. and the West sought by the Russian ruling system. Consider the following passage from a recent book by two perceptive observers of Russia and its relations with the West, Change or Decay: Russia's Dilemma and the West's Response, by Lilia Shevtsova and Andrew Wood, Carnegie Endowment for International Peace, p 214.

"A huge and prospering industry has emerged in the West. This industry includes law firms, banks, consulting firms, image makers, research centers, people in or who indirectly serve the interests of the Russian elite. Unable to modernize Russia, its elite have shown exceptional ingenuity when it comes to co-opting the West to sustain themselves and influencing Western policies."

The influence activities here do not necessarily involve anything more improper than perhaps bad or arguable political judgment on the part of the targets. But they do involve vital U.S. national security and foreign policy interests and deserve to be tracked and understood on that basis. Given the nature of the Russian side, they could evolve into more sinister activities such as money laundering, vote buying, insider trading, and "dirty tricks" in U.S. politics.

ESPIONAGE, GLOBALIZATION, AND LOYALTY

Miles D. Townes

In her important report, *Changes in Espionage*, Katherine Herbig points to the effects of globalization and the increasingly fuzzy nature of citizenship and national allegiance as an important trend in counterintelligence. Her argument is that changes in world politics have left American citizens less concrete in their allegiance, which poses a risk that those citizens entrusted with the country's secrets might seek to betray that trust.
Herbig focuses primarily on divided loyalty. For example, she draws on the PERSEREC dataset of 173 people to report that for the periods 1947 to 1979 and 1980 to 1989, "just over 20% of espionage offenders showed allegiance to a separate country or cause"; since 1990, "this proportion doubled... to 46%".[25] She argues that the changing nature of allegiance, due to globalization, will prove a difficult problem for counterintelligence in this century.

We agree generally, and here seek to amplify her point by bringing further evidence of the broader trends concerning globalization, especially with respect to citizenship and loyalty. However, we extend her analysis to dual citizens, contractors, and what we label 'diminished citizenship' -- all of which are connected to the changes wrought by globalization. In particular, we identify a troubling nexus of these concerns: the significant number of naturalized-citizen contractors involved in espionage incidents. We are concerned with the extent to which these changes are challenging traditional notions of allegiance to a single government, which allegiance is crucial to an effective personnel security system.

The point is not to impugn all such persons as unfit for national service. Rather, the point is that changes in the nature of citizenship and allegiance are pervasive. The intelligence community cannot insulate itself from these changes merely by refusing to employ naturalized citizens, dual-citizens, or contractors in sensitive positions. The changes driving those trends are much broader, affecting a wide range of even native-born American citizens.

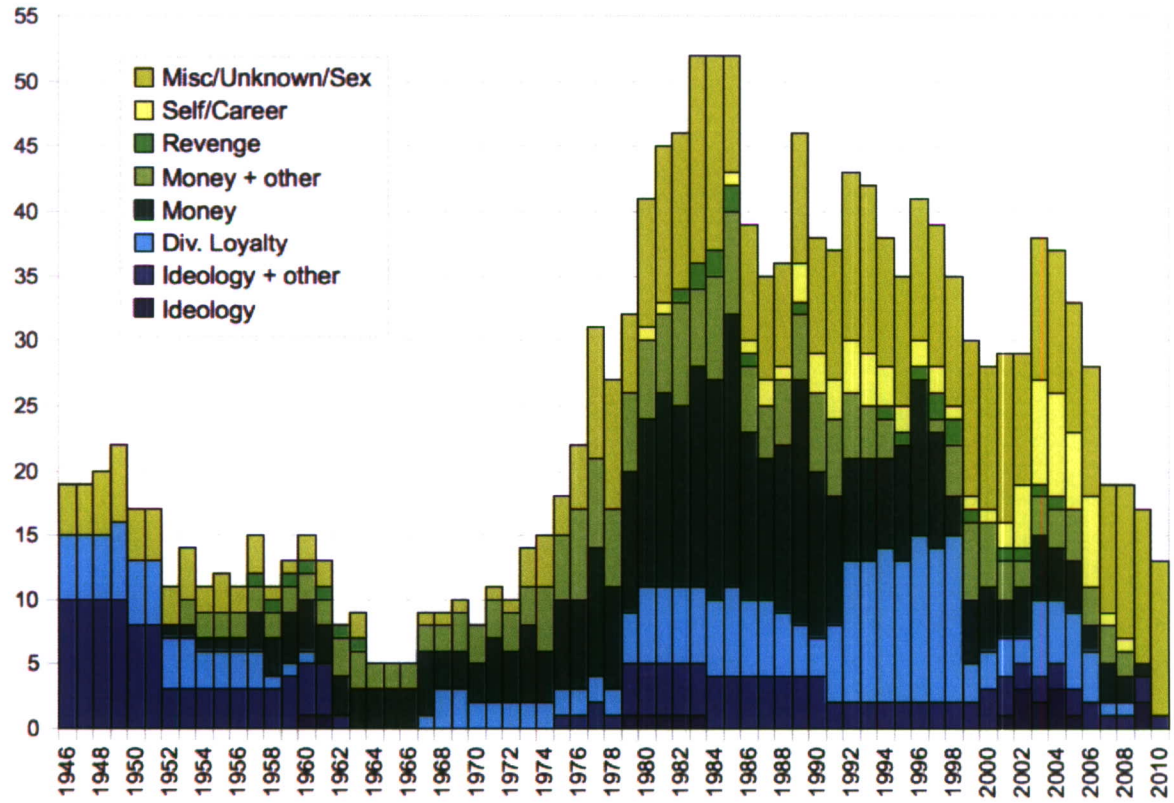
NATURALIZED CITIZENS

Herbig identifies naturalized citizens as the primary risk in a globalized world. Our data supports Herbig's conclusion; Figure 3 (reproduced from our previous report #2) shows that divided loyalty blossomed in the 1990s as a motive for espionage, and remains important into this century. Granted, many of the 1990s cases are attributable to a single spy ring, 'Red Wasp', working for Cuba, and meanwhile we coded many cases which had elements of divided loyalty as having a different dominant motivation. A number of cases we have studied in addition to Herbig's dataset comport with her understanding. Prouty, Rasool, Al-Halabi, Kim, Shu, Kuo, and Gowadia are all naturalized citizens. Leibowitz is a dual citizen. Manning's mother is British, and the two lived in Britain for several years; Manning is likely eligible for dual citizenship, as well. Of the 45 cases in which activity was initiated after 2000, at least 21 individuals' circumstances

[25] Herbig, Katherine (2008). "Changes in Espionage by Americans, 1947-2007". (Monterey, CA: Defense Personnel Security Research Center [PERSEREC]), p. 41

reflect some degree of divided loyalty; this number includes members of the Soviet 'Illegals' program, but excludes Bradley Manning. This ratio works out to 47% -- almost the same as Herbig's findings.

Figure 3
Espionage activity by motive, 1946 – 2010



Naturalized citizens are indeed an increasing share of the U.S. population. Census reports show that immigration reached a historic low by 1970, when foreign-born persons accounted for only 4.7% of the population; this trend reversed, and the 1980 Census counted 6.2 percent of the population as foreign born.[26] The percentage has continued to increase, so that the 2010 Census counted 13% of the population as foreign born, the bulk of which consists of persons originally from Latin America or Asia.[27] In absolute numbers, at present some 40 million people living in the United States hail from other countries. Of these, 17.5 million are naturalized citizens -- or 44% of foreign born, almost 6% of the total population of the United States.

---

[26] Gibson, Campbell J. & Emily Lennon (1999). "Historical Census Statistics on the Foreign-born Population". *Population Division Working Paper No. 29* (Washington, DC: U.S. Bureau of the Census). http://www.census.gov/population/www/documentation/twps0029/twps0029.html

[27] Acosta, Yesenia D & G. Patricia de la Cruz (2011). "The Foreign Born From Latin America and the Caribbean: 2010". *American Community Survey Briefs* 10-15 (Washington, DC: Census Bureau) , p. 1. http://www.census.gov/prod/2011pubs/acsbr10-15.pdf

It is easy to assume that the majority of foreign-born persons pose no interest to counter-intelligence, by virtue of their inability to obtain a security clearance for lack of U.S. citizenship. This is a mistaken assumption. A number of recent entries in our data concern individuals who either never obtained a security clearance or began their activities prior to obtaining U.S. citizenship. This is true of some members of the 'Red Wasp' ring (specifically, those who were Cuban intelligence officers), the Soviet 'Illegals' operation, and some of the other individual cases. However, these cases tended to be limited in their access to classified information. It is still difficult for foreign citizens to obtain classified information, except through native or naturalized citizen intermediaries.

Where the primary concern is those persons who spied after obtaining a security clearance, the evidence is less compelling that naturalized citizens pose a special problem. Naturalized citizens are over-represented in the data, but their cases are comparatively minor. For example, while the 'Red Wasp' ring -- which included foreign residents and naturalized citizens -- was active, Cuba's two most damaging spies in the U.S. were native citizens Montes and Myers. The aforementioned naturalized citizens -- Prouty et alia -- were also not as damaging as the native citizens active at the same time (e.g. Regan, Bergersen, Fondren).

If we discount those who were involved in economic espionage cases or otherwise had no access to classified information, the total number of relevant cases since 2000 drops to 34, of whom only ten can be counted as having divided loyalties. This ratio is only 29%, somewhat less than Herbig's estimate, but significantly higher than the percentage of foreign-born persons in the U.S. population. The reasons for that overrepresentation may be due to the true incidence rate, but it is also possible that foreign-born citizens are more likely to attract suspicion, more likely to be prosecuted, and more likely to be reported than native citizens. We can conclude that while foreign-born citizens are not the majority of the problem facing counter-intelligence, they are disproportionately represented in espionage cases.

DUAL CITIZENS

The United States is one of many countries -- including most of the developed world -- which allow dual citizenship. This policy is not the result of a specific law, but rather the 1952 Supreme Court decision in *Kawakita v. U.S.*,[28] which held that dual citizenship was long recognized in U.S. law. Although U.S. policy officially 'discourages' dual citizenship, in practice there is little the government can do to limit or restrict the practice. A 1967 Supreme Court decision, *Afroyim v. Rusk*,[29] again affirmed the Constitutionality of dual citizenship, and ruled that the U.S. government cannot strip a person of U.S. citizenship apart from voluntary renunciation.

Meanwhile, a number of other countries have loosened citizenship requirements to allow dual nationals. This is especially the case in Europe, where in some cases grandchildren of European

[28] *Kawakita v. U.S.*, 343 U.S. 717 (1952), http://uniset.ca/other/cs5/190F2d506.html; Incidentally, the argument for dual citizenship in this case was used to affirm a treason conviction for an American citizen who had worked as an interpreter at a POW camp in Japan.

[29] *Afroyim v. Rusk*, 387 US 253 (1967), http://supreme.justia.com/us/387/253/case.html

citizens are eligible for citizenship. By one estimate, 40 million Americans are eligible for citizenship with another country.[30]This does not include other sub-citizenship legal categories; for example, millions of native American citizens are considered Mexican 'nationals' by the Mexican government, by virtue of being born to Mexican-immigrant parents.[31] 'National' status in this case is not the same as citizenship, but confers special legal privileges on the individual.

We are only aware of one case of espionage in which the person was a dual-citizen in good faith (meaning not hiding his original citizenship): Leibowitz.[32] He held citizenship in both the United States and Israel, and was hired by the FBI as a linguist. It is notable that Leibowitz did not spy for either country of citizenship, but instead leaked classified information to the press.

Intelligence agencies are aware of the possibility of dual citizenship, and screen for it routinely. The FBI now includes in its employment FAQs the note, "If you are a U.S. citizen and hold dual citizenship with another country, the FBI Security Division will have to review your file to make a determination if you are eligible for employment with the FBI".[33] Presumably, dual citizenship with an adversary country is an disqualification from such employment. A person who reports dual citizenship with the US and China (which doesn't allow dual citizenship in any case) should not be hired for a sensitive post.

The problem of dual citizenship thus concerns allied access to information, and our data suggest that several allied nations have an ongoing interest in gaining access to American intelligence. Seven recent cases involve espionage on behalf of Allied countries, including Israel, Taiwan, the Phillipines, and Singapore.

DIMINISHED CITIZENSHIP

The same trends which have lead to an increase in naturalized and dual citizens -- namely, globalization -- might also lead to diminished citizenship among some native U.S. citizens. For these individuals, it is not that another specific nation competes for their allegiance, but rather that the United States no longer holds their loyalty as firmly as it might have. Some of these people will describe themselves as 'citizens of the world' -- while in fact legally very much citizens of the United States. Others simply suffer from an anomie which neglects the United States' claims on their loyalties.

One rough way to measure this phenomenon is the extent to which citizens participate in their governments. There has been a general decline in voter turnout across developed democracies, although the United States has seen increases in turnout in the last two elections. Nonetheless, the recent U.S. peak of around 63% turnout compares unfavorably with European countries,

---

[30] Abramson, A. (2008). "With US in slump, dual citizenship in EU countries attracts Americans". *Palm Beach Post* 7 June; http://www.palmbeachpost.com/localnews/content/local_news/epaper/2008/06/07/s1a_dual_citizenship_0608.html

[31] OPM (2001). "Citizenship Laws of the World". (Washington, DC: Office of Personnel Management Investigations Service); http://www.opm.gov/extra/investigate/is-01.pdf, p 133

[32] Some sources allege that Kadish was a dual citizen (also with Israel), but we consider this unverified.

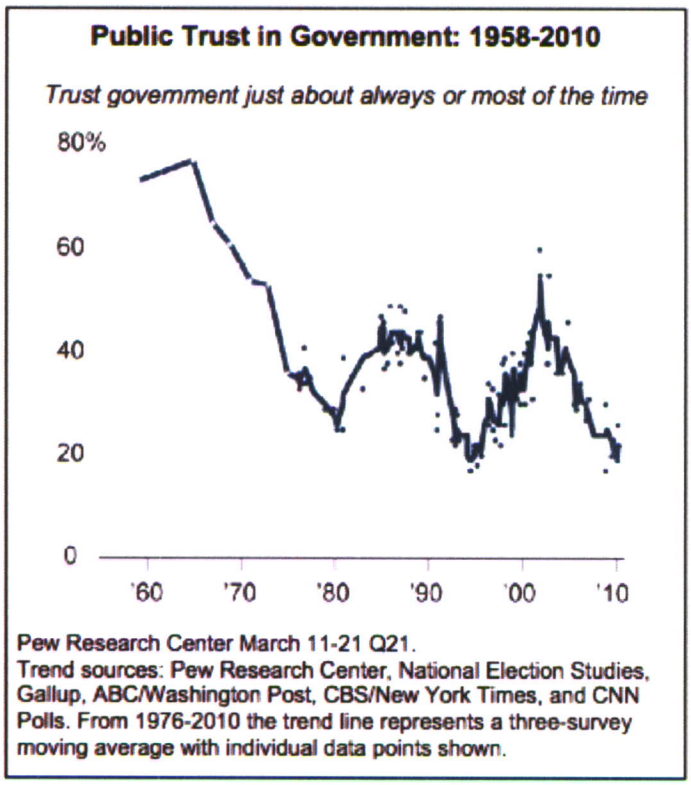
[33] FBI Careers. "FAQ". http://www.fbijobs.gov/61.asp, question 16

which are generally in the 70% range.[34] This is a rough measure, but suggests a comparative difference in terms of U.S. civic engagement.

Another indirect measure is to ask individuals to describe their trust in government. Research on trust demonstrates that people with low trust tend to be less trustworthy; while this research tends to focus on interpersonal trust, this is related to persons' trust in institutions. Thus we should be concerned when the U.S. government lacks the trust of its citizens, because that may indicate a general decrease in citizens' willingness to maintain the government's trust with respect to sensitive information.

A report from the Pew Research Center for the People & the Press shows that polls of trust in government are at unusual lows. Figure 4 shows a graph excerpted from the Pew report, which tracks the percentage of persons indicating that they trust the government 'just about always' or 'most of the time' - which is currently around 22%. Prior lows came in 1992-1995 (at 17%) and 1978-1980 (at 25%).[35]

Figure 4
Public trust in government: 1958-2010



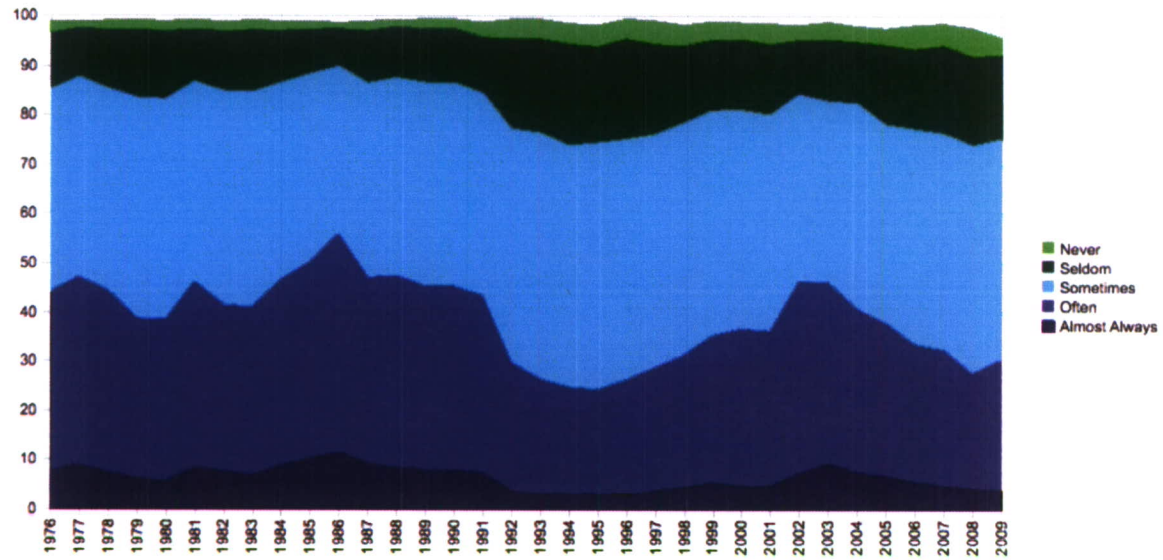
[34] McDonald, Michael. "Voter Turnout". United States Elections Project, http://elections.gmu.edu/voter_turnout.htm; see also Rosenau, James (2008). *People Count!* (NY: Paradigm), pp. 34-35

[35] Pew 2010. "The People and Their Government: DISTRUST, DISCONTENT, ANGER AND PARTISAN RANCOR". (Washington, DC: Pew Research Center for the People & the Press) April 18, 2010; http://www.people-press.org/files/legacy-pdf/606.pdf, p. 13

Similar data come from *Monitoring the Future*, a survey project of young people run by the University of Michigan, has asked each years' respondents, "how much of the time do you trust the government in Washington to do what's right?" Figure 3 shows the data in percentage form for the various response categories - Never, Seldom, Sometimes, Often, and Always. (The narrow white band at the top of the graph represents missing data.)

Figure 5 shows the most volatility between categories "Sometimes" and "Often", and we can interpolate any number of triggering events over these variations -- wars, recessions, elections, et cetera. In 1991, there was a steep drop in "Often", leading to an all-time low around 1995. Confidence in government -- measured as the sum of "Always" and "Often", then increased from 1996 to 2000. In 2002 there was a steep increase -- likely due to the attacks of September 11th. Since then confidence has been declining, with a slight uptick in 2008.

Figure 5.
Monitoring the Future responses to
"How much of the time do you trust the government in Washington to do what's right?"
by percentage, 1976 to 2009[36]



What is striking about the graph is not its variation, but rather its stability. These data generally correspond to those from the Pew report, above. Since the 1970s, when the MTF survey began, trust in government has been at generally low levels relative to its peak in the late 1950s and early 1960s, with some dynamism attributable to specific events.

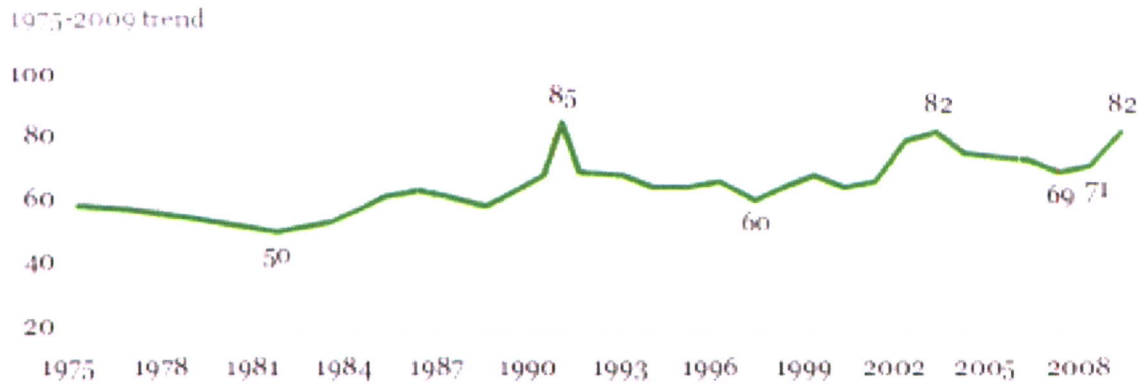
Another, more complex series of measures has been proposed by political scientist Robert Putnam. He argues that the United States has seen a decline in 'social capital'-- that is, "social connections and the attendant norms and trust" fostered by such connections -- which decline

[36] http://www.icpsr.umich.edu/icpsrweb/SAMHDA/; Data for 1990 are unavailable; we have duplicated the 1989 to produce this graph.

diminishes the democratic functions of society.[37] This concern has been taken up by the Saguaro Seminar at Harvard University, which in a 2000 report pointed out that a number of facets of civic engagement are declining: voting turnout rates, how much attention people pay to politics, campaign volunteerism, and so on. The authors conclude that "the decline in participation is troublesome for the simple reason that civic engagement is a necessary condition for wise, responsible, and effective government".[38] The disaffection of citizens also points to potential problems of trust, as an outgrowth of their connection to that government.

One caveat to the general decline or depression in trust in government is that trust in the military has been increasing since the 1980s. According to Gallup polling data,[39] trust in the military reached its low of 50% in the early part of that decade, but has climbed to 82% recently -- a point met once before in the aftermath of the September 11th, 2001, and exceeded only during the 1991 Persian Gulf War. This is shown in Figure 6. This suggests that some aspects of trust relevant to national security may have a somewhat different dynamic than trust in government more broadly.

Figure 6
Confidence in military



All of the cases in our database represent a betrayal of trust, which makes it difficult to assess the extent to which disaffection and declining trust in government play a role in those cases. In some specific cases, the individual has defended their activities by pointing to a lack of confidence in the US government -- especially in the Myers and Montes cases, where Cuba was held up as an exemplar government. This is also especially the case for the 'leakers' -- Drake, Tamm, Radack, Diaz, and Manning -- who cite a lack of confidence in the government's ability to 'do the right thing' as a reason for their behavior.

---

[37] Putnam, Robert (1995). "Tuning In, Tuning Out: The Strange Disappearance of Social Capital in America". *PS: Political Science and Politics* 28:4 (Dec.); p. 665

[38] Saguaro Seminar (2001). "Better Together", 2nd Ed. (Cambridge, MA: Kennedy School of Government, Harvard University). http://www.bettertogether.org/pdfs/bt_30_87.pdf; p. 57

[39] Gallup (2009). "American's Confidence in Military Up, Banks Down". Gallup Poll, June 14-17, 2009. http://www.gallup.com/poll/121214/americans-confidence-military-banks-down.aspx

CONTRACTORS

A separate PERSEREC report, by Kramer *et alia* (2005), discuss at length the problems of modern employment practices and the changes in the 'psychological contract' between employer and employees. They argue that the unstable nature of this new dynamic can lead to disaffection and disloyalty:

> In striving to compete in the global marketplace, American organizations more often engage in practices that some employees will experience as alienating and indicative of a lack of loyalty. More employees, lacking job security and other benefits, may become disgruntled.[40]

Kramer *et alia* avoid specifying that this is a problem for the Federal government, but the implication is clear. The Federal government is extremely reliant on contractors, who make up approximately one-fifth to one-third of the cleared workforce. Contract workers make up a significant portion of the security-cleared population. Kramer *et alia* report 2.6 million Federal employees with clearances in 1992, versus 1.9 million in 2002. Neither figure includes contractors; they report that in 2002 some half a million contractors held clearances.[41]

These figures are not definitive, however. There has been growth recently in the number of cleared-contractor population, but it is still significantly below Cold War levels. In 1985, GAO testimony reported 2.7 million federal employees with security clearances and 1.5 million contractor employees - for a total of 4.2 million cleared individuals.[42] In 1994, a GAO report showed 2.3 million federal employees and 850,000 contractors holding clearances.[43] Most recently, the Office of the Director of National Intelligence (ODNI) reported 2.7 million federal employees and 1 million contractors (and 367k "others") holding clearances in 2010.[44] By ODNI numbers, contractors account for just over one-third of cleared personnel.

Contractors often lack the specific oversight that an ordinary Federal employee might have. This is not a new problem; in 1994, Herbig wrote that the proliferation of defense contractors began in the mid-1970s; "the inevitable scattering of responsibility for enforcing security regulations among so many companies, and the inability to monitor them, allowed the Soviets and many other interested nations to make inroads in industrial espionage within the defense industry"; she also points to the growth of 'black' research programs as a source of vulnerability, which "may

---

[40] Kramer, Lisa A; Richards J. Heur, Jr.; and Kent S. Crawford (2005). Technological, Social, and Economic Trends That Are Increasing U.S. Vulnerability to Insider Espionage. Defense Personnel Security Research Center [PERSEREC] Tech. Report 05-10. p. 16

[41] Kramer, Lisa A; Richards J. Heur, Jr.; and Kent S. Crawford (2005). Technological, Social, and Economic Trends That Are Increasing U.S. Vulnerability to Insider Espionage. Defense Personnel Security Research Center [PERSEREC] Tech. Report 05-10. p. 18

[42] Thurman, Bill. "Improvements Needed in the Government's Personnel *Security Clearance* Program" Washington, DC. GAO, 16 April 1985 http://archive.gao.gov/d48t13/126710.pdf; the data reported is for 1982.

[43] GAO. "Personnel Security Investigations" (Washington, DC: General Accounting Office) GAO-NSIAD-94-135R

[44] ODNI 2010. "Annual Intelligence Authorization Act Report on Security Clearance Determinations for Fiscal Year 2010". (Washington, DC: Office of the Director of National Intelligence) https://www.fas.org/sgp/othergov/intel/clearance.pdf

increase the likelihood of industrial espionage by falling into patterns noted in earlier restricted programs, that is, of companies awarding increasing numbers of clearances while security enforcement grows lax".[45]

Indeed, some 29 cases in our total database (approx. 9%) represent betrayals by contract employees. Only the Navy (40 cases) and Army (33) accounted for a larger share from those cases for which we could identify an employer. Recent cases involving contractors include Mehalba, Chung, Mak, Nour, Gowadia, Shu, Roth, Oakley, Regan and Quintana -- accounting for 34% of the 29 cleared persons in our dataset who initiated espionage since 2000. Using Kramer's estimates, contractors account for only 1/5th the cleared population, but 1/3rd the espionage cases; using ODNI numbers, contractors are both 1/3rd of the cleared population and 1/3rd of espionage cases.

One case from our research stands out as emblematic of the potential problems posed by contractors: Almaliki Nour, also known as FNU LNU. These initials stand for 'First Name Unknown, Last Name Unknown', and his actual nationality is unknown. He arrived in the country in the late 1980s, claiming to be a refugee from Lebanon and using the name Almaliki Nour. He was naturalized as a citizen in 2000, and in 2003 was hired as a contract translator and issued a Top Secret security clearance. As a translator, he worked with combat troops in Iraq. Then in 2006, Nour was indicted for unauthorized possession of classified information - to which he pled guilty. In the course of the investigation, Nour admitted to lying about his name, date of birth, place of birth, and almost everything else about his identity.[46] His true identity is still unknown, but this case points to failings with the vetting process leading to his employment and clearance.

The government is aware of some aspects of the challenges posed by contractors. Most of these concerns are articulated in terms of the increased expense of private contractors versus federal employees; for example, an ODNI report complains that "the IC [intelligence community] finds itself in competition with its contractors for our own employees".[47] In fact, the government estimates that 70% of its intelligence budget is spent on contracts.[48] Significant and critical intelligence functions are now performed by contractors, which means those contractors also have access to important classified information. A Senate report cautioned that reliance on contractors for intelligence functions can in some circumstances lead to "strong potential for conflicts of interest",[49] and recommends reducing the number of contractor positions in the

---

[45] Herbig, Katherine (1994). "A History of Recent American Espionage". in Sarbin, T.; R. Carney; & C. Eoyang (eds.). *Citizen Espionage: Studies in Trust and Betrayal*. (Westport, CT; Praeger, 1994). p. 62-63

[46] See Herbig, Katherine L. "Changes In Espionage by Americans, 1947-2007" (Monterey, CA: Defense Personnel Security Research Center) http://www.fas.org/sgp/library/changes.pdf

[47] ODNI (2006). "The U.S. Intelligence Community's Five Year Strategic Human Capital Plan". Office of the Director of National Intelligence. http://www.odni.gov/publications/DNIHumanCapitalStrategicPlan18October2006.pdf p. 6

[48] Shorrock, Tim (2009). Spies for Hire: the secret world of intelligence outsourcing. (NY: Simon & Schuster) p . 18

[49] U.S. Senate (2007). Intelligence Authorization Act for Fiscal Year 2008: Report together with additional views. Senate Report 110-75, http://intelligence.senate.gov/11075.pdf p. 41

intelligence community. However, it appears little attention is being paid to the specific risk contractors pose to the protection of classified information.

APPLICATION TO RTISFS

The categories described above suggest possible directions in counterintelligence and personnel security. Specifically, they point to particular categories of vulnerable persons which may be used to focus and refine RTISFS or similar systems.

The basic problem for any forensic system concerned with counter-intelligence is the extremely rare occurrence of the target behavior. Identifying 1 in 10,000 dishonest people is challenging, especially if another 100 to 1000 people are behaving in ways that only appear dishonest. RTISFS solves this problem by using feedback from users and controllers to refine its interaction with identified persons, to separate out real malicious insiders from false positives -- people who are honest, but somehow misusing their access. Given that a community-wide roll out of the system, to cover all 3 million or so cleared persons, would be infeasible, it makes sense to target specific populations for the initial phases of deployment. Ideally, those populations would be slightly richer environments for espionage, so that the system can more quickly refine its approach and demonstrate its efficacy.

Our data -- and the discussion above -- point to potential populations for initial deployment of RTISFS. From our recent (post-1999) data, we have 80 individual cases. Of these, 30 are not insiders, leaving 50 insider cases. Three of the insider cases involve economic espionage unrelated to defense or intelligence information. Of those cases involving government-related information, 32 cases involved native citizens, 14 involved naturalized citizens, and one involved a dual citizen. Table 7 breaks down these 46 cases by employment (excluding the sole dual citizen) and includes expected frequencies according to Fisher's exact test. The disproportionate representation of naturalized citizen contractors is statistically significant at the 95% threshold.

Table 7.
Actual and Expected Frequencies for Native and Naturalized Citizen Insiders

Citizenship	Federal employee	Contractor
Native (expected)	28 (25)	4 (7)
Naturalized (expected)	8 (11)	6 (3)

Table 1 assumes that native and naturalized citizens are equally represented across the employee and contractor categories. There is anecdotal evidence in our data to suggest this is not the case; for example, it appears that in some cases naturalized citizens were hired somewhat hastily to meet urgent needs for translators. There may be a similar dynamic involved in high-tech companies with defense business. The bias may be due to self-selection, or a hiring bias against

naturalized citizens for Federal employment. It would be helpful to know whether naturalized citizens are over-represented as contractors, under-represented as Federal employees, both, or neither. Barring a serious underlying problem with the data, from a counter-intelligence perspective, the result remains that naturalized citizens are disproportionately represented as contractors among known cases of espionage. This suggests that a good starting point might be the examination (or re-examination) of contractors which employ significant numbers of naturalized citizens -- especially those contractors involved in defense-related technology or short-term translator programs. Again, the Nour case discussed above is emblematic of the problem.

For RTISFS deployment and usage, the advantage to these populations is not that they represent a significant population of insiders (only 13% of the 47 cases of government-related insider espionage); rather, these contexts may offer a somewhat better prospect for detection, and a somewhat diminished likelihood of false positives.

This data also suggests ways to hone or target RTISFS to the broader population of cleared individuals by identifying potential vulnerabilities. This is the goal of the Adjudicative Guidelines used by the Defense and Intelligence Communities to deny clearance to potentially risky individuals, and the guidelines are fairly comprehensive.[50] Certainly, any behavior or information which contravenes the Adjudicative Guidelines would identify a clearance-holder as a potential risk.

Criteria derived from this discussion might extend beyond the Adjudicative Guidelines to identify persons at risk for espionage; however, these cannot be used to deny those persons their clearance. These criteria would serve as a vulnerability metric, in addition to the Adjudicative Guidelines, to help counterintelligence investigators suss out potential risks. For example, persons demonstrating significant alienation -- a lack of civic participation and engagement, or personal isolation -- may be at higher risk for espionage. Persons who articulate disaffection with or lack of trust in the US government, may also be more likely to engage in espionage. The point is not to police clearance holders' speech or associations, but instead to identify persons with extreme lack of confidence in the government, or separation from community life, as they may pose a special risk to classified information. Clearance holders who bounce from contractor to contractor may be doing so for financial motives not otherwise evident. A person who answered "Seldom" or "Never" for the question, "How much time do you trust the government in Washington to do what's right?" (see Figure 3) would not seem likely to have ironclad loyalty to the government. A person alienated from ordinary community norms of allegiance and honesty may be more vulnerable to recruitment by foreign intelligence services. These are subtle vulnerabilities, but possibly an important additional tool by which to assess vulnerability among cleared personnel. Of course, there are several examples of otherwise upright citizens engaging in espionage or similar activities; a vulnerability metric based on these categories should be considered indicative, not definitive.

[50] See, for example, PERSEREC (2005). Adjudicative Guidelines for Determining Eligibility for Access to Classified Information. http://www.dhra.mil/perserec/adr/adjguidelines/adjguidframeset.htm

CONCLUSION

'Globalization' -- by almost any definition -- is changing the way many people understand their citizenship and relate to their governments. The most obvious examples come from dual and naturalized citizens -- those persons for who by definition have less than absolute loyalty to their place of birth. Indeed, such persons do pose a somewhat higher risk for espionage than native citizens.

Globalization is also amplifying other risks, as well. The increased use of contractors by government agencies, a practice derived from 'outsourcing' in the private sector, is creating significant populations of cleared persons with access to classified information under questionable oversight. Meanwhile, a parallel trend in the United States has seen the diminishment of confidence, such that relatively few persons in the country have faith in their government. This cannot help but have an erosive effect on loyalty to that government.

The United States government -- much less its counter-intelligence officials -- cannot end globalization, nor reverse these trends easily. These changes must be recognized and adjusted for, if classified information is to be protected in the coming decades.

PERSISTENT LEAKING AND THE COUNTERINTELLIGENCE RESPONSE

Miles D. Townes

One of the notable trends in our dataset is the increase in prosecution -- as espionage cases -- of persons who 'leaked' information to the press. Of 79 individuals in our dataset of recent cases (since 1999), eight persons were accused (if not charged) with transmitting classified information to the press: Manning, Sterling, Radack, Tamm, Diaz, Drake, S. Kim, and Leibowitz. "Publication" is the third most common beneficiary of betrayal of trust in this subset of our data, behind Russia and China. It is not clear whether this trend represents an increase in overall numbers of leaks, or an increase in the prosecution of such cases as espionage.

In any case, leaking must be a concern for counterintelligence personnel. On the one hand, unauthorized disclosure of classified information can in theory cause at least as much damage as compromise of the same information to a foreign power. On the other hand, leaks tend to be discrete bits of information, not ongoing compromises of classified data. Leaks also pose a set of legal and ethical issues not relevant in proper espionage cases. Leaks are a risk -- not a major risk, perhaps, but one that deserves attention.

The problem comes in policing leaks. The current approach seems to focus on punishment and deterrence, yet only one of our recent leak cases has been taken to court on espionage charges -- Drake -- and he was allowed to plead to a lesser charge. In a second case -- Leibowitz -- the defendant pled guilty. Three cases -- Sterling, Kim, Manning -- remain pending on similar charges. Though noteworthy, these cases account for a small fraction of overall leaking in the U.S. government. The current approach to leaks has not been a successful deterrent, in part because so few people are punished. A better approach would be preventative, rather than punitive, and would ensure the integrity of classified systems while providing forensic data necessary to trace leaks to their source.

WHO LEAKS?

Leaks occur from top to bottom in the Federal government; it is a pervasive problem. Some of these leaks are inadvertent. In 2005, a senior intelligence official "committed one of the biggest intelligence gaffes in recent history when she accidentally disclosed the nation's intelligence budget for that year - $44 billion. Her blooper... marked the first time since 1998 that the aggregate figure for U.S. spending on its spy agencies had been revealed...."[51] Even Presidents sometimes stumble: in 1986, President Reagan "inadvertently revealed that the NSA had intercepted Libyan communication, a secret of the highest magnitude".[52] In these cases, involving public statements, there is at least the advantage that the perpetrator is known, sparing the need for an investigation. In neither case was the leaker punished.

---

[51] Shorrock, Tim (2008). *Spies for Hire.* (NY: Simon & Schuster) p, 230.

[52] Wettering, Frank (2000). "Counterintelligence: The Broken Triad". *International Journal of Intelligence and Counterintelligence* 13

In some cases the leaks may involve information that the leaker is not aware is sensitive or classified. For example, the journalist behind a 1958 story:

> ... that US intelligence was able to monitor Soviet missile tests had no idea of the consequences of his revelation.... When the Soviets learned the Americans were monitoring their tests, they cut the advance warning time in half -- an action that forced major changes in US monitoring practices and ultimately cost millions of taxpayers dollars. The information may have been an insignificant piece of a puzzle to a reporter, but it revealed a clear picture to the Soviets.[53]

The above cases can categorized as 'mistakes'; as such, they are not our primary concern. Instead, the more problematic leaks are those done deliberately, often anonymously, with full or partial awareness of the significance of the information in question. Such leaks are the focus of this discussion.

The definitive study of leaking was published in 1986, and there is every reason to think its conclusions are still valid. In the study, which surveyed 483 former high-ranking Federal officials, some 42% admitted leaking -- "and it is reasonable to assume that the figure is, if anything, understated".[54] Of these, "nearly four out of five of the leakers identified countering false or misleading information as the reason they leaked. In their eyes, at least, they were assisting the process of getting at the truth, of helping the reporter do the job, of keeping the public informed"; other major reasons for leaking included putting something on the agenda, consolidating support from the public, or forcing action on an issue.[55] A Congressional Report in 1997 came to similar conclusions: "It has now become routine for information of the highest classification to appear in the press, most commonly as a tactical move in some intra-government policy dispute".[56] Another reporter argues that 'backgrounders' are in effect an institutionalized means for higher-level officials to leak sensitive information:

> In Washington, the same senior officials who deplore leaks and warn that they imperil national security regularly hold "backgrounders," calling in reporters to discuss policies, intelligence information and other sensitive issues with the understanding that the information can be attributed only to "administration officials" or some other similarly vague source.[57]

The author of the 1986 study concludes: "everything we have found argues that leaks as broadly defined are a routine and generally accepted part of the policymaking process.... both journalists and officials with whom we talked confirmed the view that they are a pervasive element of the

---

[53] Taylor, Stan. "Counterintelligence failures in the United States". in Johnson, Loch K., ed. (2007). *Handbook of Intelligence Studies* (NY: Routledge) p. 245

[54] Linsky, Martin (1986). *Impact: How the Press Affects Federal Policymaking* (NY: Norton), p. 172, 230

[55] Linsky, Martin (1986). Impact: How the Press Affects Federal Policymaking (NY: Norton), p. 196

[56] U.S. Congress (1997). "Report of the Commission on Protecting and Reducing Government Secrecy". (Washington, DC: Government Printing Office) Senate Document 105-2; p. A-2; http://www.gpo.gov/congress/commissions/secrecy/

[57] Wise, David (2011). "Leaks and the Law: The Story of Thomas Drake". Smithsonian, August 2011. http://www.smithsonianmag.com/history-archaeology/Leaks-and-the-Law-The-Story-of-Thomas-Drake.html

interaction".[58] A journalist speaking in the 1990s concurred: "From the point of view of a journalist, everyone in government will talk about something that they technically should not discuss. But the higher the person in government is, the more likely that seems true. At the highest levels, government officials will talk (at least on background) about almost anything to some degree".[59] This last statement is cause for concern, insofar as the higher ranks of government serve as role models for their subordinates.

Leaking is pervasive in the Federal government, and in some cases has revealed highly classified information to the general public. Yet very few people are ever punished for their leaks.

PUNISHING LEAKERS

Despite the pervasive fact of leaks as a compromise of classified information, there have been relatively few attempts to prosecute the perpetrators. The goal of this section is not to review the legal theory and statutory authority behind such cases, rather to examine the actual results of those cases in terms of their punishment of leakers and consequent deterrent effect.

The most ambitious -- and controversial -- attempts to punish leakers have been pursued under laws commonly believed to be part of the 1917 Espionage Act. In fact, the relevant provision was added to the law in 1950, and among other things states that anyone authorized to access information pertaining to national defense, who "willfully retains the same and fails to deliver it to the officer or employee of the United States entitled to receive it" should receive a sentence of up to ten years in prison.[60] This created the crime of "Willful Retention", which is the standard charge in leak cases. Although the language of the law seems straightforward, one review argues "the legislation is in many respects incomprehensible".[61] Another review elaborates:

> The relevant provisions of the espionage statutes are drafted imprecisely and are more applicable to the problem of classic espionage than to leaks of classified information by government insiders. Moreover, application of the espionage statutes to person who leaks classified information may present constitutional problems.[62]

Indeed, the Espionage Act has not fared well against leakers. There have only been three tests in court of the Espionage Act as applied to information leaks to the press, and only one has been a definite win for the government. Two of these cases occurred before 1999:

---

[58] Linsky, Martin (1986). Impact: How the Press Affects Federal Policymaking (NY: Norton), p. 196

[59] Armstrong, Scott, discussant in Theodore Sarbin, ed. (1996). *Vision 2021: Security Issues for the Next Quarter Century Proceedings* (McLean, VA: June 25-26); http://handle.dtic.mil/100.2/ADA325057, p. 82

[60] U.S. Code 18 Pt. 1, Ch. 37, sec. 793(e) ; http://www.law.cornell.edu/uscode/18/usc_sec_18_00000793----000-.html

[61] Edgar, Harold and Benno C. Schmidt (1973). "The Espionage Statutes and Publication of Defense Information" *Columbia Law Review* 73:5 (May); p. 934

[62] Ballou, E.E. and K.E. McSlarrow (1985). "Plugging the Leak: The Case for a Legislative Resolution of the Conflict between the Demands of Secrecy and the Need for an Open Government". *Virginia Law Review* 71:5 (June), pp. 805

> The first case was that of Daniel Ellsberg, who in 1971 leaked the Pentagon Papers, a secret history of the Vietnam War, to the *New York Times*. Two years later, Judge William Byrne Jr. dismissed the charges against Ellsberg due to "improper government conduct," [...]
>
> Next came the Reagan administration's prosecution of Samuel Loring Morison, a Navy intelligence analyst convicted in 1985 and sentenced to two years in prison for leaking -- to *Jane's Defence Weekly*, the British military publication -- three satellite photos of a Soviet ship under construction. After Morison was released from prison, he was pardoned by President Bill Clinton.[63]

The third case is that of Thomas Drake, indicted in 2009 for leaking information about the NSA to a reporter. The indictment against Drake asserted that he "willfully retained top-secret defense documents that he had sworn an oath to protect, sneaking them out of the intelligence agency's headquarters, at Fort Meade, Maryland, and taking them home, for the purpose of 'unauthorized disclosure'," in contravention of the Espionage Act.[64] Drake faced a prison sentence of up to 35 years, but in 2011 the government dropped its indictment and allowed him to plead to a single misdemeanor -- not derived from the Espionage Act -- for which he received no prison time.[65] The judge in the case not only rejected the government's request for a large fine, but offered a scathing rebuke of the government's behavior in the case.[66] Meanwhile, Drake was awarded the Ridenhour Prize for Truth Telling earlier in the year.[67] Although Drake was punished for his activities, and this may have some deterrent effect, the overall sense of the case is that he was more victim than villain.

The Espionage Act has only one successful prosecution against persons who leaked information to the press. In the other two cases -- Ellsberg and Drake -- the defendants have been elevated to folk heroes among those critical of government secrecy. In a fourth case, Shamai Leibowitz pled guilty to disclosing classified information to a blogger, and was sentenced to 20 months in prison[68] -- the longest sentence of any leaker charged under the Espionage Act. Meanwhile, there are three cases still pending for leaks to the press under the Espionage Act,

---

[63] Wise, David (2011). "Leaks and the Law: The Story of Thomas Drake". Smithsonian, August 2011. http://www.smithsonianmag.com/history-archaeology/Leaks-and-the-Law-The-Story-of-Thomas-Drake.html

[64] Mayer, Jane (2011). "The Secret Sharer", *New Yorker* (May 23); http://www.newyorker.com/reporting/2011/05/23/110523fa_fact_mayer?currentPage=all

[65] Nakashima, Ellen (2011). "Ex-NSA official Thomas Drake to plead guilty to misdemeanor". *Washington Post*, July 9; http://www.washingtonpost.com/national/national-security/ex-nsa-manager-has-reportedly-twice-rejected-plea-bargains-in-espionage-act-case/2011/06/09/AG89ZHNH_story.html

[66] *United States v. Thomas Drake*, "Transcript of Proceedings: Sentencing" July 15, 2011; https://www.fas.org/sgp/jud/drake/071511-transcript.pdf

[67] "Thomas Drake: 2011 Recipient of The Ridenhour Truth-Telling Prize"; http://www.ridenhour.org/recipients_03i.shtml

[68] Aftergood, Steven (2010). "Jail Sentence Imposed in Leak Case". *Secrecy News* (blog) May 25th; https://www.fas.org/blog/secrecy/2010/05/jail_leak.html

> ... including those against Pfc. Bradley Manning, a former Army intelligence analyst accused of passing State Department and military war data to the anti-secrecy Web site *WikiLeaks*; Stephen Kim, a former State Department analyst accused of leaking to a television reporter; and Jeffrey Sterling, a former CIA analyst accused of passing classified information to author and *New York Times* reporter James Risen.[69]

Whether these prosecutions will be successful is anybody's guess, but the precedent at this point suggests difficulties for the government. Also noteworthy are those cases in which the facts are materially similar, but the government chose not to pursue prosecution under the Espionage Act. Thomas Tamm was the source of leaks to the *New York Times* concerning warrantless wiretap surveillance; in fact, these leaks were the origin of the investigation that led to Thomas Drake's indictment for unrelated activity. However, earlier this year the Justice Department dropped its investigation of Tamm.[70] Likewise, Jesselyn Radack leaked information about the government's treatment of an accused terrorist to Newsweek; she was fired from her job and placed under criminal investigation, but no charges were ever brought.[71] When the identity of covert CIA officer Valerie Plame was leaked to the press, some argued that the Espionage Act could be brought to bear against the person responsible.[72] Despite a lengthy investigation, nobody was ever charged with the actual leak in the Plame case.

One of the difficulties in prosecuting leaks cases is that the investigations often run through the journalists who publish the information. Bradley Manning was identified as the source of major leaks of government information because a person who offered him confidentiality as a reporter in fact contacted the FBI. Otherwise, journalists typically resist these investigations, and some have been willing to go to jail to protect sources. Judith Miller of the *New York Times*, for example, spent twelve weeks incarcerated, rather than reveal her source for the Plame leak. After being subpoenaed to testify about his interactions with Jeffrey Sterling, *Times* reporter James Risen stated, "I am going to fight this subpoena... I will always protect my sources, and I think this is a fight about the First Amendment and the freedom of the press".[73] This is despite the fact that the Supreme Court ruled in *Branzburg v. Hayes* that the First Amendment does not protect reporters from being required to testify in federal courts.[74]

---

[69] Nakashima, Ellen (2011). "Case Narrows Against Thomas Drake, ex-NSA manager accused of mishandling classified files". *Washington Post*, June 8; http://www.washingtonpost.com/national/national-security/case-against-ex-nsa-manager-accused-of-mishandling-classified-files-narrows/2011/06/07/AGk3ZZMH_story_1.html

[70] Memmott, Mark (2022). "Justice Drops Probe of Leaker Who Exposed Bush-Era Wiretapping". *National Public Radio: The Two-Way* (blog), April 26. https://www.npr.org/blogs/thetwo-way/2011/04/26/135735752/report-justice-drops-probe-of-leaker-who-exposed-bush-era-wiretapping

[71] Horton, Scott (2010). "Justice's Vendetta Against a Whistleblower: Six Questions for Jesselyn Radack". Harper's Magazine: No Comment (blog), Feb. 23; http://harpers.org/archive/2010/02/hbc-90006592

[72] e.g. Dean, John (2003). "The Bush Administration Adopts a Worse-than-Nixonian Tactic: The Deadly Serious Crime of Naming CIA operatives". Findlaw (blog), Aug. 15; http://writ.news.findlaw.com/dean/20030815.html. "Scooter" Libby was convicted of perjury stemming from the investigation, not of any crime specific to the leak.

[73] Savage, Charlie (2011). "Subpoena Issued to Writer in C.I.A.-Iran Leak Case". New York Times, May 24 2011; https://www.nytimes.com/2011/05/25/us/25subpoena.html

[74] Toobin, Jeffrey (2006). "Name that Source; why are the courts leaning on journalists?" *The New Yorker*, 16 January, p. 30

The Espionage Act, while it may be the appropriate legal instrument by which to address leaks of classified information, is in fact unwieldy in practice. Punishment of leaks is uncertain and inconsistent, which substantially limits the desired deterrent effect. One argument might be that the laws should be tightened, but this would be difficult given First Amendment protections and probable resistance from the media. Moreover, it is not clear how much tighter the laws need to be in order to override the strong -- if perhaps misguided -- moral sensibility of persons like Diaz, Radack, and Tamm. All three were lawyers for the government, and thus acutely aware of the potential for punishment inherent in their activities, yet they chose to leak nonetheless.

PREVENTING LEAKS

Given the difficulties in prosecuting leaks, a better approach to protecting classified information is to prevent leaks in the first place. Protection means monitoring the use of classified information and ensuring users are accountable for their behavior. Careful monitoring of such information could, in many instances, provide forensic information which helps avoid lengthy and expensive investigations -- especially in those cases where the journalists involved refuse to be cooperative. Such preventive measures are likely cheaper in the long run than lengthy investigations and trials, which offer uncertain outcomes anyway.

A comprehensive forensic system, such as RTISFS, would protect best against the most damaging kinds of leaks -- those who compromise information on an ongoing basis, or compromise significant amounts of information at once. These leakers are most likely to have their behavior flagged by the RTISFS system, and thus require further investigation. Rather than being subject to the whims of the press, RTISFS allows controllers to focus on the big risks first.

RTISFS also allows investigators to determine the severity of the compromise, and act accordingly. This follows from the typology introduced in our previous report, which describes a range of insiders from fundamentally honest to utterly malicious. For many trivial would-be leaks cases, a clear policy which punishes offenders with loss of clearance privileges will be sufficient deterrent -- so long as they have reason to believe they will be caught.

The problem of leaks is rightly the concern of counterintelligence. The current approach makes "example" cases of leaks, treating them as seriously as espionage -- but the vast majority of leakers will leak with impunity, no matter how severe punishment the few caught persons receive. A better approach would be preventive approach based on good protection of classified information, integrated forensic analysis, and a flexible typology of offense.

Task 2 – Responses to Changes in User Behavior Online

REFINING THE DESIGN OF THE INTERROGATORY PROCESS

Rebecca Givner-Forbes

LYING AND HUMAN NATURE: RELEVANCE FOR IDENTIFYING DECEPTION

In general, human beings are bad lie detectors. The famed deception researcher Paul Ekman posits an evolutionary reason. For most of human development, says Ekman, we lived in close-knit kinship groups with almost no privacy, in which the reputational consequences of lying could have a permanent and even life-threatening impact. Modern societies, on the other hand, provide almost limitless opportunities for lying and one can always move to escape the reputational cost of lying. For this reason, lying has increased, but our ability to detect liars has not yet developed.[75]

Most liars, while motivated by the opportunities and advantages lying can afford in our modern, achievement-oriented society, have similarly not developed the skills to lie without leaking signs of deception. While human beings have not yet developed the intuitive ability to detect these cues, they can be detected by specially trained individuals or by automated techniques. Most people do not accurately identify these signs, however, including those in job positions that would seem to require the skill, like law enforcement professionals. [76]

Whether Ekman's evolutionary explanation is convincing or not, research supports his general conclusion that people are bad at detecting deception.[77] The research has consequences for insider detection. Existing counterintelligence depends to some degree on the monitoring by supervisors and coworkers. Nigel West points out that "few spies are caught as a result of the 'vigilance of colleagues' or routine security screening." Rather, most spies are caught when identified by active sources or defectors.[78]

Another major component of counterintelligence are routine personnel screenings. As West suggests, these also leave something to be desired. Typically, they use polygraph tests and processes to vet lists of foreign contacts. While polygraphs are more effective than unassisted humans, they are not full proof spy-detectors. A report from the Defense Personnel Security Research Center points out that at least six Americans have managed to spy while passing personnel security vetting procedures and maintaining their security clearances.[79]

---

[75] Ekman, P (1996). Why Don't We Catch Liars, *Social Research* 63:3, pp. 801-817.

[76] Memon, A., A. Vrij, & R. Bull (2003). Psychology and Law: Truthfulness, Accuracy and Credibility. (West Sussex, England: Wiley), p. 29-33.

[77] *e.g.*, Newman, M, J. Pennebaker, D. Berry, and J. Richards (2003). Lying Words: Predicting Deception from Linguistic Styles. *Personality and Social Psychology Bulletin*, 29, pp.665-675.

[78] West, N (2007). "Cold War Intelligence Defectors," in Loch K. Johnson (ed.), *Handbook of Intelligence Studies* (Abingdon, UK; Routledge) pp. 229-230

[79] Herbig, K. and M. Wiskoff (2002). Espionage Against the United States by American Citizens 1947-2001. Monterey, CA: Defense Personnel Security Research Center, p. xiii.

People may be even more incapable of detecting deception from colleagues in the workplace than they are at detecting it from strangers. Ekman points out that "involvement in a relationship can lead to confidence in one's ability to detect deception…and such confidence may itself make one more vulnerable." [80]We have a relationship with our coworkers, we think that we know them. Yet many people, empirically speaking, have been surprised to discover something about a coworker. When the stakes are extremely high, the insider is more likely to focus attention on successfully deceiving colleagues, and his colleagues may be less likely to suspect him of the worst kinds of crimes.

Ekman provides an example of the unconscious desire to collude in a lie in a high-stakes situation. During the September 1938 meeting between then-British Prime Minister Nevelle Chamberlain and Adolph Hitler, Chamberlain was so eager to avoid war, and had placed himself in such a risky political position by neglecting to prepare for confrontation, that he staunchly defended Hitler as man of his word who would not carry out further acts of aggression in Europe. "Chamberlain was not unique," writes Ekman. "The targets of lies, often unwittingly, collusively want to believe the liar…One is nearly always better off in the short run to cooperate with the lie, even if that means the consequences tomorrow will be even worse." [81]

Many individuals do not closely monitor their coworkers at work (and current hiring freezes within the government mean that employees will be busier than ever with their own workloads). Further the hierarchical nature of many government agencies mean individuals may be likely to defer to a supervisor, especially when it comes to such grave matters as fingering potential traitors to the nation. Such a supervisor may have had a hand in hiring the insider at issue, and therefore may fall unwitting victim to the Chamberlain effect. Most people do not think of themselves as easily duped, and so the longer they work with a person, the less likely they are to suddenly become suspicious of him. Such suspicious would require an admission that, until this point, they have been tricked.

An insider detection process that supplements personnel security screening and does not rely on the chance provocation of suspicion in colleagues in between such screenings could help close the gap in effective insider detection. The purpose of this paper is not to critique traditional methods of identifying suspected government insiders or to suggest such methods should be displaced. Rather, this paper describes a process that can supplement traditional methods and procedures and help individuals responsible for counterintelligence deploy them efficiently and objectively.

---

[80] Ekman (1996).

[81] Ekman (1996).

PROSPECTS AND LIMITATIONS OF LINGUISTIC ANALYSIS FOR RTISFS

Many studies have identified linguistic metrics that change with deception. Deception produces emotional responses and cognitive stress, and this changes how people use language.[82] This is especially true with "high-stakes" deception, where life or liberty is at stake.[83] These linguistic metrics can be identified and analyzed to detect deceit.

Deception research describes three general theories as to why lying alters language. First, there is the "emotions perspective," which holds that deceivers feel guilt over lying and/or fear of being caught.[84] This emotional response manifests itself in comments that reflect a negative emotional state, such as aversion, negation, anxiety, and anger.[85] Fear of being caught, or distaste with lying, manifests itself in the use of general, vague, or indirect responses. A lack of specificity provides ambiguity and reduces the chance of being definitively caught in a lie.[86] The desire to distance oneself from the lie manifests in a reduction in first person pronouns and a general tendency for liars not to refer specifically to themselves.[87] The desire to distance one's self from the lie results in the use of more tentative words and fewer words that connote certainty.[88]

Second, the "cognitive effort" theory holds that fabricating lies is cognitively difficult.[89] If the liar does not have time to prepare a story, the lie may lack detail and be nonspecific.[90] A liar who has the opportunity to rehearse or polish an answer will take the time to do so. The liar will hesitate before beginning his response to a question, and will also make more hesitations during speech.[91]

The "control perspective" theory is consistent with the above two theories and comes to similar conclusions. It states that deceivers who do not want to be caught– whether because of an emotional or social aversion to being outed as liars, or because of more serious consequences like criminal prosecution – will exhibit certain linguistic characteristics consistent with the effort

---

[82] Newman and Pennebaker.

[83] Memon, Vrij, and Bull, p. 18. Vrij A. *et al* (2007). Cues to Deception and Ability to Detect Lies as a Function of Police Interview Styles, *Law & Human Behavior*, 31, 499-518.

[84] Larcker, D. and A. Zakolyukina (2010). Detecting Deceptive Discussions in Conference Calls. Stanford GSB Research Paper No. 2060, Rock Center for Corporate Governance Working Paper No. 83, p. 7.

[85] Gupta, S (2007). Modeling Deception Detection in Text. Thesis Submitted to the School of Computing at Queen's University, Kingston, Ontario, Canada. p. 10.

[86] Larcker and Zakolyukina, p. 8.

[87] This particular deceit cue is the most widely supported in the literature reviewed for this paper. *E.g.*, Hancock, J. et al (2008). On Lying and Being Lied To: A Linguistic Analysis of Deception in Computer-Mediated Communication. *Discourse Processes*, 45, pp. 1-23. Newman & Pennebaker. Larcker & Zakolyukina, p. 7.

[88] Adams, S. and J. Jarvis (2006). Indicators of Veracity and Deception: An Analysis of Written Statements Made to Police. *Speech, Language, and the Law* 13:1, pp. 1-22, Bond, G. & A. Lee (2005). Language of Lies in Prison: Linguistic Classification of Prisoners' Truthful and Deceptive Natural Language. *Applied Cognitive Psychology* 19:3, pp. 313-329.

[89] Newman and Pennebaker.

[90] Larcker and Zakolyukina, p. 8.

[91] Memon, Vrij, & Bull, p. 19.

to exert control over the listener and avoid being revealed. Subconsciously, they will avoid self-references, which tend to produce an emotional response. They instead distance themselves from the lie with third-person pronouns so they can tell it more dispassionately.[92] They use a greater number of unique words to achieve "lexical diversity" in order to sound more convincing.[93] When people tell the truth, they tend to repeat statements, leading to fewer unique words.

Consistent with the emotional perspective theory, control perspective theory holds that liars will use less detail and more general terms if they do not have the opportunity to prepare a lie.[94] However, when liars have the ability to rehearse a lie, the control perspective theorists claim that lies can be more detailed and more specific in an effort to be convincing.[95] Because of the increased detail, their responses will be longer.[96] Also, in contrast with the emotions perspective theory, liars who are engaging in "impression management" to be likable and convincing to their audience may express fewer negative emotions.[97]

Other studies do not postulate theories for why lying causes linguistic change, but they support the view that certain linguistic characteristics, or metrics, correlate with deception. A study of lies in asynchronous computer communications show that liars in this particular medium produce more words than truth-tellers, probably because computer mediated communications provide an opportunity to review, edit, and save text. It provide the liar with the opportunity to rehearse or polish a statement and look back at it later, reducing the chance that the liar will contradict himself.[98]

Liars commit more errors and mistakes in their speech.[99] They use fewer "exclusive" words than truth-tellers – such as "but," "except," and "without," because making fine distinctions between what is in a given category and what is not requires an extra level of cognitive complexity.[100] As one study's authors put it, "exclusive words create fine distinctions in one's story that could later be disproved." [101]

---

[92] Hancock, J. et al. Zhou, L. et al (2004). Automated Linguistics Based Cues for Detecting Deception in Text-Based Asynchronous Computer-Mediated Communication: An Empirical Investigation. *Group Decision and Negotiation*, 13:1, pp. 81-106.

[93] Larcker and Zakolyukina, p. 9.

[94] Larcker and Zakolyukina, p. 8.

[95] Burgoon, J., Blair, J., Qin. T., & Nunamaker, J. (2003). Detecting Deception Through Linguistic Analysis. *Lecture Notes In Computer Science: Proceedings of Intelligence and Security Informatics*, 2665, pp. 91-101.

[96] Larcker and Zakolyukina, p. 9.

[97] Memon, Vrij & Bull, p. 12.

[98] Hancock, J. et al.

[99] Newman and Pennebaker.

[100] Newman and Pennebaker.

[101] Newman and Pennebaker.

Studies of deception use either manual coding or automated coding to classify language along the features to be analyzed. In manual coding, a person reads transcripts of responses and classifies words and phrases along the dimensions being measured for deception. Other studies make use of automated psychosocial dictionaries that identify and classify words, phrases, or sentences automatically. The most well-known of such dictionaries is the linguistic inventory and word count (LIWC) developed by James Pennebaker.[102]

LIWC categorizes speech across 72 different dimensions. It was originally developed as a tool to assist mental health researchers in identifying certain mental illnesses that, like deception, manifest measurable linguistic characteristics.[103] When making a model to measure deceit, only a small number of the 72 dimensions are analyzed. Models using LIWC to predict deception in academic studies typically measure the use of first-person pronouns, third-person pronouns, exclusive words, negative-emotion words, and action-verbs.[104] For example, LIWC will parse a text and put words like "worried," "concerned," and "annoyed" into the "negative-emotion" category.[105] Deception models using these factors demonstrate accuracy rates ranging from 61-69 percent in detecting deceit.[106]

Manual coding models require the use of a trained analyst who reads and intelligently classifies text across multiple dimensions. The best known among these are Criteria Based Content Analysis (CBCA) and Reality Monitoring (RM). CBCA was originally developed to assess the credibility of alleged victims of child sex abuse, but has demonstrated promise to detect deception in a broad variety of other contexts since its development.[107] Both CBCA and RM work off the basic assumption that a person recalling actual events will provide different kinds of information than a person fabricating a memory. This is known as the Undeutsch hypothesis.[108]

Studies employing CBCA analyze a statement across 14-19 metrics.[109] The most commonly included metrics are the following: logical structure of the statement; contextual embeddings (references to time and space); descriptions of interactions; reproduction of speech (quoting others); accounts of subjective mental state; spontaneous corrections; and admitting lack of memory.[110]

---

[102] http://www.liwc.net/

[103] Pennebaker, J. Mehl, M. and Niederhoffer, K (2003). Psychological Aspects of Natural Language Use: Our Worlds, Our Selves. *Annual Review of Psychology*, 54, pp. 547-77.

[104] Gupta, p. 13.

[105] http://www.liwc.net/descriptiontable1.php

[106] Gupta, 13. Newman & Pennebaker.

[107] Vrij (2007), 501.

[108] Vrij (2007), p. 501.

[109] Fuller, C (2008). High-Stakes, Real-World Deception: An Examination of the Process of Deception and Deception Detection Using Linguistic-Based Cues. Thesis Submitted to Oklahoma State University, p. 14.

[110] Vrij (2007), p. 501.

Studies vary widely in reported accuracy of CBCA to detect deception, from 55-90 percent.[111] One reason for this is CBCA's reliance on trained coders to provide ratings to statements. Human coders vary in terms of the ratings they apply to the same statements. Efforts to automate coding or provide better training to make coders' ratings more uniform are likely to improve accuracy rates. The lesson from this is that the fewer coders an organization uses to code interrogatory responses and the more they work together to standardize coding, the more accurate CBCA is likely to be.[112] There are no formal rules for determining how these criteria are weighted, and how many of them or in what frequency they must be present in order to determine whether a statement is truthful or deceptive.[113]

RM uses the following to identify fabricated recollections:

- In deceptive recollections, less sensory information (memories of real experiences are likely to contain more details of smell, taste, touch, as well as visual and auditory details)
- In deceptive recollections, less contextual information (memories of real experiences contain more spatial details and details about how other people and objects were situated in relation to each other, e.g., he stood behind me and temporal information - details about the timing of events - first this happened, then this).
- In deceptive recollections, more cognitive operations like reasoning (e.g., I must have had my coat on, as it was very cold that night).[114]

RM has advantages over CBCA. It allows for more standardization among raters, which makes it easier to use and probably increases accuracy. Some studies comparing the two show that RM is more accurate than CBCA; others show it has similar accuracy.[115] The basic linguistic analysis tool proposed here can not make use of RM or CBCA because their focus on recalling memories, but approaches discussed later could employ these techniques.

Linguistic analysis has limitations. In the controlled studies analyzed for this paper, success rates for identifying a statement as truthful or deceptive averaged in the mid-60 percent. However, aspects of the specific context of interrogatory responses indicate potentially higher accuracy rates. First, the stakes for insiders are much higher than in the kind of controlled laboratory experiments in which much deception research is conducted. High-stakes lying, one can hypothesize, produces more tension and therefore more of the linguistic characteristics that indicate deception.[116] RTISFS can also capture metrics that theorists have postulated exist in deception, but have not been included in deception detection models used in studies. These include recording the time taken to compose a response as well as a user's efforts to edit and

---

[111] Fuller, p. 16.

[112] Vrij, A., K. Edward, K. Roberts, and R. Bull. (2000). Detecting Deceit Via Analysis of Verbal and Nonverbal Behavior. *Journal of Nonverbal Behavior*, 24:4, 239-263.

[113] Fuller, p. 17.

[114] Vrij (2007), p. 502.

[115] Vrij (2007), p. 502.

[116] Fuller, pp. 30-31. Memon, Vrij, & Bull, p. 18.

polish it. Linguistic analysis results can also be analyzed in tandem with other technical factors such as whether a user's linguistic profile correlates with deception at the same time that their technical use profile shows abnormalities – too many files accessed or copied, for instance. Lastly, linguistic deception model can improve itself over time by capturing metrics across several users in a single job category over a long time period, as well as over a span of time for a single individual.

LINGUISTIC METRICS ANALYZED AUTOMATICALLY

Not all the dimensions/metrics shown to indicate deception in the above studies are appropriate to the unique context in which RTISFS is intended to be used. For example, because RTISFS will mostly query a user on what the user is doing at that moment, and not his perceived recollection of an event, CBCA and RM will not be immediately useful for analyzing the responses to many interrogatories.

The model we propose in Table 8 depends on collecting and categorizing data along the metrics noted below for automated analysis. This entire model is referred to as LAC because it is the component of RTISFS.

Table 8
Proposed model for dimensions for automated analysis

DIMENSION ANALYZED	HOW COLLECTED	WHY COLLECTED
Hesitation, efforts to construct a message	RTISFS measures elapsed time between when respondent begins typing and when respondent submits interrogatory; controlled for number of words.	Detects efforts to structure a convincing response. Increased hesitation positively correlates with deceit.
Efforts to polish a response or errors made	RTISFS measures characters backspaced or deleted while composing responses	Deceivers more likely to rewrite answers; make errors.
Saving an interrogatory response	RTISFS measures attempts to copy response text, print or print screen interrogatory response, and total elapsed time from when the interrogatory response is entered in its final form to when it is submitted (e.g. when a person may be taking time to manually write down what they have written).	A deceiver, especially a trained insider, would benefit from saving responses so that they could be recalled consistently later.
Restarting a computer when an interrogatory is triggered	RTISFS would collect this information.	A user who does not want to respond to an interrogatory may restart the computer and attempt to find the information another way. A user may also be exploring the limits of the system to discover what events trigger interrogatories.
First person pronouns	Automatically identified by LIWC	Negatively correlated with deception
Third person pronouns	Automatically identified by LIWC	Positively correlated with deception

DIMENSION ANALYZED	HOW COLLECTED	WHY COLLECTED
Negative emotion words	Automatically identified by LIWC	May be positively correlated with deception or negatively correlated
Certainty words	Automatically identified by LIWC	Negatively correlated with deception
Tentative words	Automatically identified by LIWC	Positively correlated with deception
Distinction words (e.g. except, but, without)	Automatically identified by LIWC	Negatively correlated with deception
Word count	Automatically counted by LIWC	Liars who have time to compose a lie are likely to have longer responses
Lexical diversity	Automatically counted by LIWC	Positively correlated with deception

Research examined in connection with this study did not provide data that would assist in assigning different weights to different metrics above to reflect the strength of the correlation with deception. This may vary somewhat according to the type of experiment and the context. For example, some experiments lend themselves to responses from subjects that would produce more "negative-emotion" words. A study that had respondents lie about an emotionally-charged topic would produce more negative-emotion words in deceptive responses. On the other hand, in the workplace context, negative-emotion words may not be viewed as acceptable. For now, each metric should be given equal weight, but this is an area where RTISFS could be empirically be "tuned." RTISFS should also be customizable so that a CIC can exclude any metrics that appear to be not relevant or useful.

BUILDING USER PROFILES TO MEASURE AUTOMATICALLY-COLLECTED METRICS

For any given metric for a given type of interrogatory, the system would build baselines or averages by collecting data from a class of users, as well as from a single user over time. RTISFS would identify outlier responses whose intensity or frequency could trigger an interrogatory The "sensitivity" of the system would be set by the policies and priorities of the deploying organization and could be adjusted over time.

As an example shown in Table 9, take a simple system that had two classes of interrogatories and two trigger events defined, each categorized by the event that triggered them and the nature of the interrogatory question itself:

Table 9
Two classes of trigger events

TRIGGER EVENT	BAIT QUESTION	OPEN-ENDED QUERY
1. Accessing a certain number of files within a 30-minute period.	Interrogatory 1: what concern would your supervisor express over this activity and why?	Interrogatory 2: What is the reason for your current activity.

TRIGGER EVENT	BAIT QUESTION	OPEN-ENDED QUERY
2. Attempting to print or print-screen a sensitive document	Interrogatory 3a could use the same bait question for a user who had already triggered even t1 as above, or Interrogatory 3b could be premised on trigger event 1 having been committed	Interrogatory 4a could be independent of having committed trigger event 1 or Interrogatory 4b could be premised on trigger event I having been triggered.

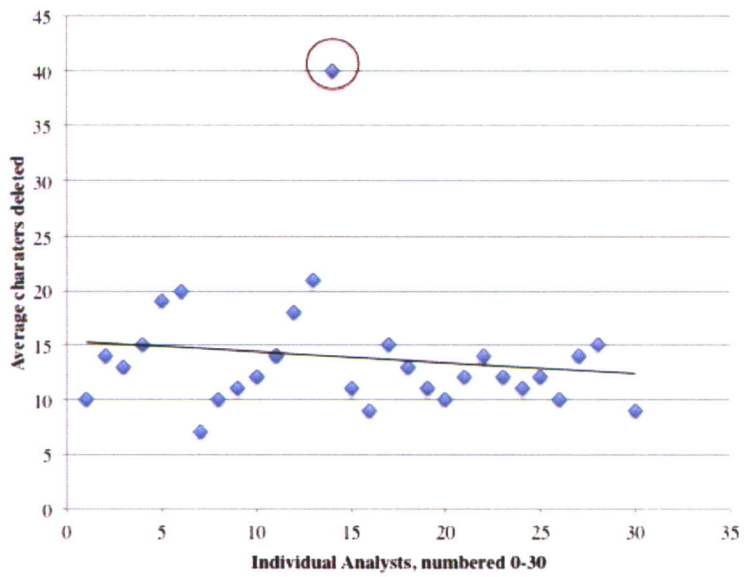
Because the vast the majority of users will providing honest responses, as seen from the earlier paper on the incidence rates of trust violations, the system can spot for deviations from this baseline. The two charts below represent the responses to interrogatories .

As an example, say you have a group of 30 users in a similar class of profession (e.g., a particular group of Department of Defense analysts). Their interrogatory responses are analyzed across the metrics that correlate with deception, using the average of the 30 as a baseline. Say you have one user (user 14) for whom, compared to other similar users, shows significant deviations (in whatever manner the system is calibrated to define "significant") across three metrics potentially indicative of deception. The first two of these indicators are shown in the accompanying tables, which simply represent how one user can be shown to deviate from a baseline for a particular metric.

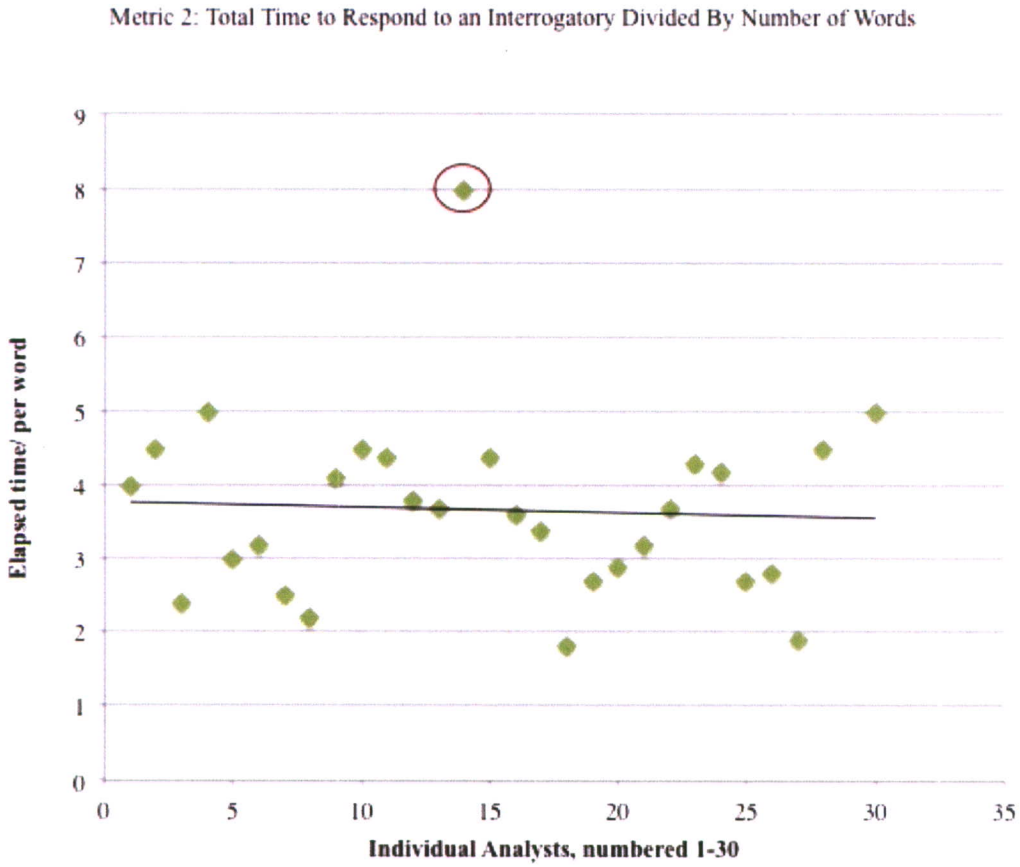
The first chart (Metric 1 below) is the number of characters deleted while formulating an interrogatory response, suggestive of excessive polishing or tailoring of an answer that will "pass muster." This is correlated with deception.

Two Metrics Demonstrating Single User's Efforts to Polish
Interrogatory Responses Compared to Other Users

Metric 1: Average Number of Characters Deleted



The second metric (Metric 2 below) for which this user's interrogatory responses depart from the norm is in the time it took for the user to respond to an interrogatory (divided by the number of words, suggesting careful word selection). This is also positively correlated with deception.



User 14's interrogatory responses also deviate from the baseline in terms of pronoun use, with less first-person pronouns. (This is a third metic, the table of which is not shown here).

Such a result is not by itself definitive. It could simply be that user 14 does not have English as a native language and is working on spelling and syntax while dealing with the interrogatory – hence more deletions and a longer period of time spent responding. Perhaps user 14 is a native Mandarin Chinese speaker, whose pronoun use in English is influenced by the fact that, in Chinese, fewer first person pronouns are used. Some linguistic features can carry over from a native to an acquired language, as a result of a person translating literally in their head from their native tongue.

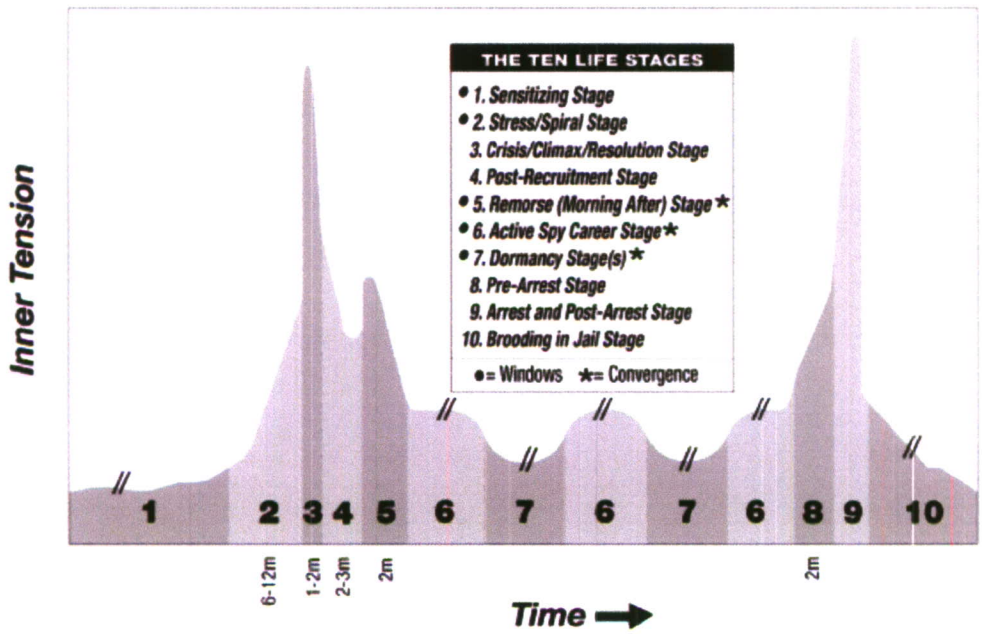
The example demonstrates that the collection of statistics can assist in establishing the probability of such an outlier, but it will, in any case, necessitate take a closer look at the particular user.

The development of a model of deception through the analysis of user text input in a unique context over a period of time is likely to have greater predictive power than that in many of the above-discussed studies, which often use generic models based on general theories of deception to try to predict deception in a specific context. Developing a model of deception that is based on deviations from an average or baseline from many responses by similar users in similar contexts is likely to have strong predictive power.[117] Over time, the model will become more accurate, and it will be increasingly clear which metrics are the most fruitful indicators of deception in a particular context.

LOOKING FOR ABNORMALITIES BASED ON A USER'S PROFILE

The functionality that captures a user's own data over the course of time and looks for deviations from his own norm is a response to earlier case studies that many insiders turn disloyal only after they join an organization. Working with people who have turned disloyal shows that shows they pass through various emotional and psychological stages, some of which are characterized by periods of intense stress. This stress can increase fluctuations along the linguistic dimensions analyzed noted above. The chart (Figure 7 below), taken from research by Dr. David Charney, MD, and noted in our previous report #2, illustrates this.[118]

Figure 7
Ten life stages of an insider spy



[117] Newman and Pennebaker.

[118] Charney, D. (Fall 2010), True Psychology of the Insider Spy, *Intelligence Journal of U.S. Intelligence Studies*.

A user profile should change at stage 2. Assuming that RTISFS has monitored the user for a sufficiently long period of time to establish a baseline during stage 1, the additional stress should manifest itself sufficiently to demonstrate detectable linguistic changes in stages 2 and 3. This comes about as close as logically possible to an early DARPA vision of identifying malicious insiders *before* they did committed any violations.

Tracking interrogatory responses across the metrics associated for deception also helps establish a baseline for a particular user who may have "turned," or who may begin giving deceptive responses partway through his career. Referring back to the example in the section above discussing "user 14", a look at his profile reveals that the deviations his responses show across the three metrics identified could be explainable by the fact that English is his second language. One way to confirm this is to look at those metrics for the entirety of his career, should such material be available. If the user has never before shown such a large deviation across these metrics, than that suggests that language issues are not causing the deviation.

An organization deploying RTISFS will have to set the sensitivity of the system based on its priorities and resources. A system that flags too many users in an agency that does not have the resources to investigate all of them is not useful. Adjusting the sensitivity of the system, (i.e. the user profile variations required to flag a user for further scrutiny), will be needed to control the false positive rate. In view of the relative infrequency of malicious users, the detection probability will have to be addressed through red team efforts experimentally.

DEVELOPING INTERROGATORIES

The development of interrogatory content is one area in which we found little concrete research applicable to this format. As such, different kinds of interrogatories querying a user regarding the reasons for his activities or asking him to describe them should be deployed in a testing environment to determine which are most effective.

Honesty testing in the workplace, which gained some popularity during the 1980's, consisted of paper-and-pencil quizzes given to employees that purported to assess their honesty or other desirable character traits. They fell out of favor in part because they were being administered and evaluated in an ad hoc fashion, with evaluators providing little evidence of any particular analytical methodology.[119] They consisted of questions like "do you ever think about committing a burglary?" Or "would you say that you are too honest to steal?" These tests did not hold up to scrutiny, and do not provide a source of useful us.[120]

Research into effective interviewing and interrogation techniques may be somewhat applicable, although such techniques often depend in part on the back-and-forth nature of an interrogation, and take into account the face-to-face questioning on the psychological reactions and conduct of a suspect, including the suspect's nonverbal reactions. One finding that may be applicable from

[119] Moore, R. and R. Stewart (1989). Evaluating Employee Integrity: Moral and Methodological Problems. *Employee Responsibilities and Rights Journal*, 2:3, p. 208.

[120] Moore and Stewart, p. 212.

interrogation research is the finding that open-ended questions produce more accurate data for verbal deception analysis than accusative questions.[121]

Examples of open-ended interrogatories include:

- What is your purpose for accessing this file?
- Please describe the project for which you are conducting this search?
- What will do with this document after you print it?

These open-ended questions provide less structure for a response than more accusatory questions, allowing more freedom for the linguistic dimensions reflecting deception to manifest.

One can design additional interrogatories based on the deception theories discussed above. Because these theories hold that the linguistic features of deceptive responses appear because the deception provokes emotional or cognitive stress, we recommend questions that are specifically designed to provoke stress in a deceiver. For example:

- Accessing this file drive is unusual for individuals with your job classification. Why is this necessary to your work?
- Most people cannot read more than twenty pages of text in an hour, but you have accessed much more than that in a shorter period of time. How many of the documents you have opened in the last hour did you not read in their entirety, and why?
- What concern would your supervisor express about what you are doing right now?

Such "bait" questions are intended to produce a type of heightened emotional response in both truth tellers and deceivers, but because of the unique cognitive and emotional stress placed on deceivers, their responses to such questions manifest linguistic changes along the dimensions or metrics noted.[122]

Another class of questions would suggest to a user that his responses could be verifiable, or exposed as deceitful, in an offline investigative process. These interrogatories ask for the names of individuals associated with the job task related to the triggering event. Deceitful individuals would be more likely to give vague answers to attempt to avoid being caught out. An individual may also try to avoid responding to such an interrogatory by backing out of the triggering event or restarting their computer. Example interrogatories include:

- Who assigned the task associated with this action and what were the specific instructions that they gave to you?
- You are trying to access data regarding (subject). For what task do you need to view this data, who else is working on the task with you, and what would those individuals say about the relevance of this data to that task?

---

[121] Vrij (2007).

[122] Discussion with Dr. Sharon Smith, Falls Church, Virginia, August 29, 2011.

To develop and select proper interrogatories, a controlled testing process is necessary. Such a process would consist of deploying various kinds of interrogatories in a testing environment, with some respondents instructed to lie about what they are doing and others told to tell the truth. The testing environment will have to be designed so that incentives produce some heightened emotional response in the liars.

ACTIVITIES MONITORED AND TRIGGER EVENT SELECTION

Most of this paper is focused on the use of interrogatories, but it may also be useful to discuss the activities that may be monitored by RTISFS and the user events that trigger interrogatories.

The following describe basic user actions on a computer system, along with possible malicious insider activity that could be connected with the user action. Each of these may trigger interrogatories, depending upon the sensitivity settings of the system.

- **User Action**: Perform a search within certain parameters
  - **RTISFS** records the search, compares parameters to searches performed by others in the job category or to a job position profile for this user.
  - **Potential Insider Activity**: Searching for documents unnecessary to job task, or exploring/ testing the system to see what activity will trigger interrogatories.

- **User Action**: Creating and sending "test" files
  - **RTISFS** notes sending of attachments without text
  - **Potential Insider Activity**: Testing the system to see if emails with attachments provoke a response.

- **User Action**: Opening multiple, unrelated file directories, conducting a certain number of searches within a limited period of time, or viewing a greater-than-normal number of files in a certain period of time.
  - **RTISFS** records heightened frequency of user activity
  - **Potential Insider Activity**: Seeking potentially valuable information for non-job-related use.
- **User Action**: Retaining data by copying, printing, print-screening, attaching files to emails.
  - **RTISFS** records all actions above as efforts to retain data.
  - **Potential Insider Activity**: Attempting to create copies of data that can be removed from the facility.

- **User Action**: Attempting to change file descriptions, saving files with formats or names inconsistent with their contents, deleting contents of a "downloads" file, changing user logs.
  - **RTISFS** records the above and flags as attempts to cover system activity
  - **Potential Insider Activity**: Covering tracks or "hiding" information within the system to make it difficult to retrieve.

- **User Action**: Logging on throughout the night or at other irregular hours.
  - **RTISFS** records dates and times of user activity.

- **Potential Insider Activity**: Extracurricular insider activities that cannot be conducted easily during normal work hours.

Interrogatories should also be triggered by mundane events, or at certain time intervals, to produce sufficient amounts of text to establish useful baselines for both individuals and interrogatory classes. This second class of triggering events is also important to gather data from individuals who are possible malicious insiders, but whose insider activities do not require that they commit the more serious triggering events.

An organization will not want RTISFS to introduce too much friction into the work process. Too many interrogatories will also increase employee frustration, which may provoke emotional responses or refusal to take interrogatory inquiries seriously. A major focus of the testing phase be to establish appropriate trigger events for interrogatories, content of interrogatories, and deciding how many interrogatories can be deployed in a given time period without negatively affecting productivity.

MAINTAINING RTISFS EFFECTIVENESS OVER TIME

The linguistic deception detection techniques are based on open-source research. A malicious insider who realizes that his responses are being coded and analyzed will be able to research how to provide honest-seeming responses. Operational aspects of RTISFS should kept in the background and features added to it to provide genuinely useful user assistance. Given the limited focus of Phase I, we have not attempted to explore possible value-added by RTISFS to user and organizational performance.

Even if a malicious insider suspects his interrogatory responses are designed to gather data for counterintelligence purposes, he will not know exactly what metrics are being analyzed, or even if his responses will be analyzed at all until someone suspects him. He will be unable to anticipate the collection of certain metrics, and metrics can be changed over time.
Also, well-meaning employees could disrupt the system's proper functioning by discussing answers to interrogatories with one another. If employees standardize their answers, or inform one another when they mention each other in their interrogatory responses, this would negatively effect the proper functioning of the system. It could also help the malicious insider "game" the system by ensuring his responses are consistent with those of other employees. These are operationally important issues that are likely to be addressed except in specific organizational contests.

Efficient employees, or those who become frustrated typing replies to interrogatories, may endeavor to keep a stock of interrogatory response templates in order to quickly respond to interrogatories. This would disrupt the integrity of the interrogatory process. An obvious protection would be to not allow pasting responses into interrogatory text boxes

DEALING WITH PSYCHOPATHS AND SOCIOPATHS[123]

Some malicious insiders may be categorized as psychopaths or sociopaths. Such individuals have no conscience and are expert liars, and so will not exhibit the same indicators of deceit as people with less aptitude at lying and some conscience. For the purposes of this discussion, we will not distinguish between the psychopaths and sociopaths; they both exhibit the relevant traits of adept lying and a lack of conscience relevant to this analysis.

One prominent researcher on sociopaths has posited that 1 in every 25 individuals may be sociopaths.[124] Such individuals may be especially prone to become malicious insiders if employed in sensitive government positions because of their total focus on the self and their lack of loyalty to other people or institutions.

In the case of socio- or psychopaths, rather than look for deceit along the metrics described above, other metrics typical of these conditions must be considered. The metrics described here are derived from the experience of FBI agents interrogating psychopathic criminals. Some of these can be automatically detected and categorized by LIWC.

These include a greater use of first-person singular pronouns, reflecting the extreme focus on the self. The benefit of using a model that detects deviations from an established average or baseline is that it can detect a deviation in either direction: a higher-than-average use of first-person pronouns typical of psychopaths and a lower-than-average use of first-person pronouns typical of ordinary deceivers. The key is deviation from a mean.

A psychopath is likely to think himself to be smarter than the system and attempt to interact with whoever is analyzing his responses. He may attempt this interaction through sarcastic or glib remarks. While LIWC is not yet sufficiently sophisticated to identify sarcasm, attempts to communicate with the person behind the interrogatories will exhibit use of second person pronouns and also question marks. The following provide hypothetical examples of interrogatory responses from a psychopath (P) and a non-psychopath reflecting this dynamic:

- Interrogatory: What could happen if the information you are about to access fell into the wrong hands?
    - Non-P: My colleagues could be captured and killed.
    - P: Isn't this fancy system supposed to stop that from happening?

---

[123] The research contained in this subsection is all derived from a discussion and email exchange with and an unpublished paper by Dr. Sharon Smith, a Special Agent for the Federal Bureau of Investigation for 25 years, and an instructor at the FBI's Behavioral Science Unit with expertise in psychopaths. Dr. Smith's review of profiles of insiders collected for this project led her to conclude that some exhibit profiles consistent with psychopathy, and that such individuals may require a different detection approach.

[124] Stout, M (2005). The Sociopath Next Door. Crown Archetype.

- Interrogatory: How many of the 25 files you have accessed in the past hour have you not read in their entirety?
    - Non-P B: All of them. This data are so poorly organized I can't find what I need.
    - P: How many files can YOU read an hour?

A psychopath is also likely to blame others whenever he feels that his actions may be suspicious. He may also view the system as a tool he can use to cast suspicion on his coworkers. While certain interrogatories may request names of other people, a user who names (and blames) coworkers in response to other kinds of interrogatories that do not ask for a name may be exhibiting this characteristic:

- Interrogatory: Why are you copying the selected text discussing (subject).
    - Non-P: To paste it into an executive summary for *our* (subject) Project (the non-P views himself as part of a team, and so is more likely to own the project by using the pronoun "our," rather than describing the work as something he was told to do. This is also a first person plural pronoun, negatively correlated with both deceit and psychopathy).
    - P: Because Bob told me to. He didn't give me a reason and I don't see why he needs this information.

Some psychopathy metrics require a human analyst, however. If a user's profile is flagged and a human analyst reads interrogatory responses, the analyst may look for the following additional traits:

- Not taking serious interrogatories seriously (see the above example of the psychopath's non-serious response to the serious query: What could happen if the information you are accessing fell into the wrong hands?).
- Glib or sarcastic comments.
- Putting down others, for example:
    - Interrogatory: It is unusual for someone in your position to perform this function. Why are you doing this?
        - Non-P: I have been asked to help out another team working on this project.
        - P: It's only unusual because no one else here does there job correctly.
- Using hyperbole reflecting an inflated sense of self and the importance of one's own work, for example:
    - Interrogatory: Why are you doing _____ on this system?
        - Non-P: For...[legitimate descriptive purpose].
        - P: Because I'm trying to prevent the next September 11th !

The problem of psychopaths and sociopaths in a system – as well as society – has yet to be thoroughly explored. Further psychological research into such individuals as well as insights from the criminal justice community, particularly law enforcement and other professionals with experience interrogating and interviewing sociopaths and psychopaths, can help build more complete interrogatory response profiles.

AVENUES OF FUTURE DEVELOPMENT

Red Teaming

Deploying RTISFS in a test environment to produce initial baselines along the dimensions analyzed will be needed before it can be implemented in an operational environment. A Red Team can begin without knowledge of how the system works or what metrics it gathers for analysis, just as an employee would. Relevant data to gather during a Red Team phase includes:

- Does the Red Team guess that LIWC or other kinds of content analysis are being performed on the text of their interrogatories?
- If so, do they figure out how to "beat" this analysis?
- Do they form accurate hypotheses regarding what triggers interrogatories?
- Do they figure out ways to access data that should trigger interrogatories, but does not?
- Does a Blue Team detect this manipulation? If so, how? If not, could they have if additional metrics were collected?

A Red Team would consist of individuals with knowledge of intelligence training in certain key countries, individuals with technical knowledge that may enable them to circumvent RTISFS, psychologists or others who can represent the traditional insider psychology, and others playing the role of other classes of malicious insiders outside the traditional malicious-upon-entry trained spy or the turned intelligence asset: e.g, the Bradley Manning type of insider.

Such a Red Team process would produce valuable insights in how to help safeguard the effectiveness of the RTISFS system.

Other Sources of Text for Analysis

The system described in this paper proposes the use of linguistic analysis of only one source of text input: responses to interrogatories. It may also be possible to deploy the same kind of analytical process to other sources of text produced in the workplace, such as employee emails. Input collected and analyzed from employee emails could supplement RTISFS as a source of more free-form text output from subjects. The following describes one successful use of email analysis along the same metrics discussed in this paper to detect deceit:

> An interesting pilot study exploring what can be achieved along these lines was recently National Laboratories and New Mexico Tech. The investigation of deception detection in e-mail used the Enron e-mail corpus, a publicly available collection of approximately five hundred thousand e-mails exchanged between Enron employees and others over a three-

> year period. Both message content and e-mail metadata were analyzed. For the message content portion of the study a very simple "bag of words" model was used, so that the message was considered to be simply a set of works, and all other syntactic and semantic structure was ignored. The deception model employed was also quite simple. It was assumed that individuals engaged in deceptive informal communication exhibit reduced usage of first-person pronouns and exclusive words and increased usage of negative emotion and action words. Analysis consisted of building very large network representations of the message content for the entire Enron e-mail corpus, with messages linked to key words from the four classes of words hypothesized to be relevant for deception. Automated analysis of this network successfully identified both deceptive messages and individuals who were particularly prone to engaging in deception (as independently verified via court transcripts and other information sources).[125]

This demonstrates one successful use of the kind of model described in this paper on asynchronous computer-mediated text communication - in this case, email. Responses to interrogatories are another variety of this same kind of communication, defined by a lack of face-to-face interaction and synchronicity (dialogue/two-way chat).

In addition to the LIWC metrics discussed in this paper and used on the Enron corpus, CBCA and RM also show potential as linguistic analysis tools to detect deceit. They are more challenging and costly to deploy because they require human coders. While they analyze deceit in a person's recollection of an alleged memory, distinguishing a real recalled event or experience from a fabricated one through the use of different linguistic structures, this format can find application in some investigations.

CBCA and RM could be used in some higher level of analysis for a suspected insider; perhaps as an intermediary step between the basic RTISFS described above and more invasive or costly measures such as a FISA warrant or a polygraph. The system could include a feature that requests a long-form response to queries prompting a recollection. For example, foreign contact screening procedures are an important tool in counterespionage. If a CIC became suspicious of a users connection with a particular foreign contact, a message could be sent to the person asking him, for example, to describe in detail the first and last interaction he had with that contact. CBCA and/or RM could be run against the response for signs of deception. These analysis techniques may be able to determine if the recollections of these interactions are fabricated or real.

Increasing Automation

A system that detects deviations from the norm for a particular user across several metrics in response to a particular interrogatory could automatically compare it to other user's metrics for this type of interrogatory over a period of time. Even a malicious user is unlikely to need to be deceptive most of the time, and the baseline obtained from averaging his responses may produce

---

[125] Gosler, J (2007). Counterintelligence: Too Narrowly Practices, in Jennifer E. Sims and Burton Beber, eds. Vaults, Mirrors, and Masks: Rediscovering U.S. Counterintelligence. (Washington, DC: Georgetown University Press), pp. 188-89.

an "honest" baseline. It must be recognized that automation carries with it the risk of subtle errors compounding each other. Hence the need to incorporate human judgment to the greatest degree consistent with overall system effectiveness in detecting deceit.

The paradox is when a system such as RTISFS operates for long periods and finds nothing when, in fact there may have been nothing to uncover. To circumvent this, a user population monitored by RTISFS can be divided into subsets, and system sensitivity adjusted to collect more positives for offline analysis and preliminary investigation.

Adding a Biometric Dimension to RTISFS

The Chinese used to force rice powder into the mouths of people suspected of deceit. If the rice powder was dry when the suspect spit it out, he was judged to be a liar.[126] This technique represents an early, if crude, attempt to detect a biometric indicator of deceit: a dry mouth. Humans may be poor lie detectors, but this anecdote suggests they have long seen promise in detecting the body's physiological responses to telling a lie.

Polygraph tests are a more modern biometric lie detector. They typically measure sweating in the hands, blood pressure, and respiration. Studies to assess success rates of polygraphs in detecting deception at between 59 – 87 percent.[127] The problem is that polygraphs are disruptive, and certain people, especially sociopaths/psychopaths and trained spies may be able to beat them more easily than ordinary people.

However, adding some kind of biometric dimension to the detection deceit could increase its accuracy. Biometric indicators collected surreptitiously would mitigate both the problem of individuals who can beat polygraphs as well as the disruption that too frequent polygraph testing would cause to a workplace. Biometric data captured could be factored into RTISFS to detect physiological arousal in a user while he is responding to interrogatories or performing suspicious or unusual actions on an information system. Even a sociopath or psychopath who has no emotional response to lying because of a lack of a conscience may still experience signs of arousal associated with a fear of being caught or "duper's delight" – excitement at fooling others.[128]

Recent studies of thermal warming demonstrate that a rise in temperature under the thin skin around the eyes can be detected. Some researchers claim that these findings show the "potential for application in remote and rapid security screening, without the need for skilled staff or physical contact." [129]An employees mouse or keyboard could also measure sweating hands, like a polygraph. A highly sensitive microphone could detect respiration rate. A person normally rests their hands on a bar or the bottom of keyboard while typing; sensors could record pulse. The incorporation of cameras and microphones into computers and portable devices makes a wide class of add-ons in employer provided productivity tools potentially useful. Since all devices

---

[126] Memon, Vrij, and Bull, p. 20.

[127] Memon, Vrij, and Bull p. 26.

[128] Memon, Vrij, and Bull, p. 19.

[129] Pavlidis, J., Eberhardt, N. and Levine, J (2002). Seeing Through the Face of Deception. *Nature*, 415.

derive their functionality from software, it can be introduced remotely through the use of all devices on information networks.

JUDGING SUCCESS

The success of RTISFS can be measured in terms of how it closes the gap between the number of insiders and spies that are identified and the number believed to be at work. The accompanying paper, "Incidents of Violations of Trust" suggests an order of magnitude of between 10^{-3} and 10^{-4} representing trust violators (a rough proxy for malicious insiders) for a national security or defense organization for any period of time. Five to ten years is the proper scope of time to judge the success of this system in those terms. Such a period is adequate to build accurate baselines and user profiles and allow CICs to become adept at its use.

PROACTIVE COUNTERINTELLIGENCE FOR THE INFORMATION AGE

Miles D. Townes

We argue that adopting a broad definition of counterintelligence -- for the policies and agencies encompassing that definition -- is the best way to approach the challenges of protecting classified information in the Information Age. The Cold War focus on adversary nations and their intelligence agencies no longer is adequate in light of the diverse threats and risks now facing the intelligence community.

Instead, counterintelligence should be understood at its broadest, to include not only active counterespionage but also preventative measures: personnel security, physical security, and information security, including active information protection systems like RTISFS. Where counterintelligence was focused in the Cold War on the malicious individuals who were already spying, it should now focus equally on protecting information *ab initio*.

If intelligence is finding information, counterintelligence must go beyond finding the finders.

COLD WAR DEFINITIONS

The Cold War intelligence community was almost entirely -- and rightly -- concerned with the Soviet Union and its allies; the result is that intelligence and counterintelligence came to be defined with implicit reference to the Soviet Union. Because the Soviet Union operated according to an intelligence strategy which was largely dependent on human intelligence -- spies -- counterintelligence evolved in response a preoccupation with catching spies. The clearest official statement of this understanding came in Executive Order 12036 (1978), which

> describes CI as both "information gathered" and "activities conducted", the purpose of which is to "protect against espionage and other clandestine intelligence activities, sabotage, international terrorist activities or assassinations conducted for or on behalf of foreign powers, organizations or persons, *but not including personnel, physical, document or communications security*".[130]

That summary is derived from a 1979 symposium which attempted to forecast the intelligence environment of the 1980s. Note that the distinction between document and communications security in particular is dated, and can be usefully combined under 'information security'.

The above quote appears in a paper titled, "What is Counter-intelligence?", by Arthur Zuehlke, Jr. A discussant for the paper offered his own definition:

> ... counterintelligence is the art of of examining the entire spectrum of enemy intelligence activities in the light of enemy intelligence strategy, our own and allied intelligence and counterintelligence services, for the purpose of devising better means of advancing our policies

[130] Zuehlke, Arthur A. Jr. "What is counterintelligence?" in Godson, Roy, ed. (1980). *Intelligence Requirements for the 1980's: Counterintelligence.* (New Brunswick, USA; Transaction Books), p. 24, emphasis original.

> and protecting our nation from espionage, subversion, disinformation and deception and adverse military and political and economic actions. The government must regard counterintelligence as the essential tempering ingredient that will harden the weapons to dull the Soviet strategic sword -- the KGB.[131]

Although this could be interpreted as more expansive than EO 12036, it is clear by the final sentence that the author is in fact focused on the Soviet Union. On the other hand, a second discussant offered this view:

> Although I recognize the definition in EO 12036 (presumably for bureaucratic, regulatory and oversight reasons) as excluding personnel, physical, document, or communications security, I do not share [Zuehlke's] enthusiasm for their exclusion from the "true function" of CI , but rather believe they are essential ingredients of any systems approach to looking at the total of CI systems.[132]

Indeed, this view seems to be shared by several of the participants in the symposium. In a later paper in the symposium, on counterintelligence organization, one author writes: "Counter-intelligence is concerned with the protection of information from those who are seeking to obtain it. Counterintelligence consists of three components: personnel security, physical security, and counter-espionage".[133] Another author writes along similar lines, "Distinctions are often drawn among operational security, physical security, and security of personnel. Yet these disciplines are so intermeshed that a failure in one always jeopardizes the other two".[134] Despite the government's extant commitments, the debate is never resolved in the volume in favor of one view over another. As the editor of the volume said in his introduction, "It is not easy to define counterintelligence. Its practitioners themselves disagree about the meaning of concept".[135] Such disagreement has persisted into the present.

CHANGES AND CHALLENGES IN THE INFORMATION AGE

Cold War counterintelligence was focused on the malicious individuals serving enemy interests, and very specifically those working for the Soviet Union and its intelligence apparatus. The dynamic facing the United States in the information age is very different: there is no single focus, no single strategy against which we can shape our response. In this environment, the Cold War focus on malicious individuals faces a number of shortcomings, and counterespionage is less viable as a comprehensive response to the threat.

---

[131] Miler, Newton (1980). discussant in Godson, Roy, ed. (1980). *Intelligence Requirements for the 1980's: Counterintelligence.* (New Brunswick, USA; Transaction Books), p. 41-42
[132] Thompson, Edmund. R, Maj. Gen. discussant in Godson, Roy, ed. (1980). *Intelligence Requirements for the 1980's: Counterintelligence.* (New Brunswick, USA; Transaction Books), p. 45
[133] Smith, Norman L. (1980). "Counterintelligence Organization". Godson, Roy, ed. *Intelligence Requirements for the 1980's: Counterintelligence.* (New Brunswick, USA; Transaction Books), p. 214
[134] Pratt, Donovan (1980). "Counterintelligence Organization". Godson, Roy, ed. *Intelligence Requirements for the 1980's: Counterintelligence.* (New Brunswick, USA; Transaction Books), p. 229
135 Godson, Roy, ed. (1980). Intelligence Requirements for the 1980's: Counterintelligence. (New Brunswick, USA; Transaction Books), p.1

Foremost, the task of counterintelligence can no longer focused primarily on a single government, a convenience of the Cold War. A recent PERSEREC document says, "Individuals in both government and industry in almost 100 countries were involved in legal and illegal efforts to collect intelligence in the United States during 2004".[136] While serving as National Counterintelligence Executive, Michelle Van Cleave reported that as many as 140 nations and 35 terrorist organizations are engaged intelligence activity against the United States.[137] Even using a lower estimate, this still represents a staggering number of governments, languages, cultures, and strategies against which counterintelligence officers must organize and act. It is unlikely that the counterintelligence agencies will be able to recruit sufficient personnel qualified to deal with this broad array of threats while still focused on counterespionage.

Second, there is the possibility that misuse of classified information may not involve foreign organizations, but rather domestic organizations and activist groups. Consider that Diaz sent his classified information to an activist group; it was the judgment of a single member of that group (a lawyer who might have faced disbarment otherwise) that the group should report the breach, and not use the information. Clients for illicit classified information may include domestic governmental organizations, as in the LA County Sheriff's spy ring; granted, the Sheriff's office is not as severe a threat as the Soviet Union, but classified information was seriously compromised in that case just the same -- and by sworn law enforcement personnel. For reasons discussed elsewhere in this report, it is not safe to assume that information which leaks into 'friendly' organizations will be kept in confidence by those organizations.

A third problem is the potential for 'penetration' of foreign organizations not affiliated with foreign governments. Despite being a preoccupation of the Cold War era, US counterintelligence was never especially successful in penetrating the Soviet apparatus (judging from available evidence). The KGB alone employed some half million personnel.[138] Many terrorist organizations rely on a small cadre of trusted individuals; these are notoriously difficult to penetrate. Advocacy organizations like Amnesty International also tend to be cloistered and protective of their sources; Wikileaks in particular operates along very insular, personality-driven lines.[139] The 'offensive' approach -- which relies on compromises of foreign intelligence services -- is less viable when no such service exists. Nor does such an approach afford any protection against leakers and 'information should be free' advocates, who decide independently and anonymously to compromise classified information.

A fourth problem is that narrowly-defined concepts of counterintelligence have lead to an emphasis on counterespionage over other, intimately related domains -- per Executive Order

---

[136] PERSEREC (2010). "Counter-intelligence" in Adjudicative Desk Reference (online resource) http://www.dhra.mil/perserec/adr/counterintelligence/counterintelligenceframeset.htm

[137] Van Cleave, Michelle (2005). "Prepared Statement of Michelle Van Cleave" in House Committee on the Judiciary, Subcommittee on Immigration, Border Security, and Claims. "Sources and Methods of Foreign Nationals Engaged in Economic and Military Espionage". 109th Cong., 1st. sess., 15 September 2005. http://www.gpo.gov/fdsys/pkg/CHRG-109hhrg23433/pdf/CHRG-109hhrg23433.pdf , p. 11

[138] Pike, John (1997). "KGB/Sources and Methods". Federation of American Scientists Intelligence Resource Program; https://www.fas.org/irp/world/russia/kgb/su0515.htm

[139] Poulsen, Kevin and Kim Zetter (2010). "Unpublished Iraq War Logs Trigger Internal Wikileaks Revolt". Wired: Threat Level (blog; September 27, 2010) http://www.wired.com/threatlevel/2010/09/wikileaks-revolt/

12036. As one expert explains, the mindset embodied in that order has led to the deprecation of 'security' as an important task of counterintelligence:

> Counterintelligence officers, especially at the Central Intelligence Agency (CIA), tend to dismiss the protection of secrets as ``merely'' security. Indeed, in the counterintelligence profession "security" officers are looked down on as poor cousins who have to deal with safe closings and employee thefts rather than the exciting business of catching spies. This hubris has resulted in a split throughout both the federal government and the private sector which has resulted in two bureaucracies: ``security'' and ``counterintelligence.'' Yet, physical and personnel security are actually major components of counterintelligence.[140]

The same expert summarizes the consequences of this split as follows:

> United States counterintelligence is alive but not well. Its triad of three essential functions is: protecting secrets, frustrating attempts by foreign intelligence services to acquire those secrets, and catching Americans who spy for those foreign intelligence services. The first of these functions is in effect broken, that is, not being performed.[141]

The attention paid to foreign intelligence services comes at the expense of attention to security -- information security, personnel security, and physical security. In an era when information can be beamed wirelessly from cell phone to anywhere on the planet, the divide between 'security' and 'counterintelligence' is a contrived distinction and a dysfunctional policy.

COUNTERINTELLIGENCE SINCE THE COLD WAR

Despite the obvious challenges of the information age, many observers cling to the Cold War understanding of counterintelligence. In a recent volume on counterintelligence, the general understanding from the many authors is that counterintelligence should continue to focus on spies and foreign intelligence operations. One author argues "the need for counterespionage to play a larger role in the nation's counterintelligence strategy".[142] Arguably, counterespionage has in fact been the dominant approach the counterintelligence -- by the definition of EO 12333 -- but the same author admits, "We face the reality that persons convicted for espionage represent a tiny percentage of the American citizens and foreign nationals who operate as either the agents or intelligence officers in the United States".[143]A recent author writes that "the objective of every counter-intelligence organization is to identify, penetrate, and then control or neutralize its

---

[140] Wettering, Frank (2000). "Counterintelligence: The Broken Triad". International Journal of Intelligence and Counterintelligence 13, p. 266

[141] Wettering, Frank (2000). "Counterintelligence: The Broken Triad". *International Journal of Intelligence and Counterintelligence* 13, p. 265

[142] Wallace, Robert (2009). "A Time for Counter-Espionage", in Jennifer Sims and Burton Gerber, Vaults, Masks and Mirrors: Rediscovering U.S. Counterintelligence (Washington, DC: Georgetown U. Press), p. 101. (NB: We assume the book was already in publication when EO 13470 was issued, and thus the editors and authors were not able to address its implications for their understanding of counterintelligence.)

[143] Wallace, Robert (2009). "A Time for Counter-Espionage", in Jennifer Sims and Burton Gerber, Vaults, Masks and Mirrors: Rediscovering U.S. Counterintelligence (Washington, DC: Georgetown U. Press), p. 104

adversary".[144] Another: "counterintelligence is the process of countering the hostile intelligence activities of other states or foreign entities".[145] These are all worthy activities, and important components of an integrated, comprehensive counterintelligence strategy -- but they are not themselves such a strategy.

Recently the Federal government has begun to recognize the changing nature of the intelligence threat, and the need for a broad response. Executive Order 12036 was superseded in 1981 by Executive Order 12333, which retained the narrow definition of counterintelligence and adversarial focus.[146] EO 12333 remained unchanged as the guiding authority in the intelligence community for the next two decades, and was reflected in the National Counterintelligence Strategy of 2007, which again excluded "personnel, physical, document or communications security" from the scope of counterintelligence.[147] Despite revisions in 2003 and 2004, not until EO 13470 in 2008 was the definition of counterintelligence amended to account for changes in the world since the end of the Cold War.[148] According to the 2008 amended version of EO 12333, counterintelligence is now defined as:

> information gathered and activities conducted to identify, deceive, exploit, disrupt, or protect against espionage, other intelligence activities, sabotage, or assassinations conducted for or on behalf of foreign powers, organizations, or persons, or their agents, or international terrorist organizations or activities.[149]

This definition does acknowledges the scope of counterintelligence challenge in the information age. In particular, it no longer excludes personnel security, physical security, and information security from consideration as legitimate activities. However, it does retain an emphasis on foreign or hostile intelligence activities. This is a halfway point, at best.

The 2009 National Counterintelligence Strategy shows that policy makers are not yet thinking comprehensively enough. The document -- like its predecessors -- is preoccupied with foreign spies. At only point does it discuss security as a function of counter-intelligence. From the section titled, "Protecting the Integrity of the U.S. Intelligence System" comes this:

> The U.S. intelligence system must provide reliable information to the U.S. government and its allies. The integrity and reliability of this system – the people, the structure, the information systems, and the information they hold – depend on our ability to keep it free from penetration or influence. In pursuit of this objective, the counterintelligence community will work closely with our colleagues in security, acquisition, information assurance, and other relevant specialties

---

[144] West, Nigel (2007). "Cold War Intelligence Defectors". in Johnson, Loch K. ed. Handbook of Intelligence Studies (NY; Routledge), p. 229.
[145] Taylor, Stan (2007). "Counterintelligence failures in the United States". Johnson, Loch K., ed. Handbook of Intelligence Studies (NY: Routledge), p. 237
[146] President (1981). "Executive Order 12333 -- U.S. intelligence activities" Federal Register 46, p. 59941 http://www.archives.gov/federal-register/codification/executive-order/12333.html
[147] Gosler, James R. (2009). "Counterintelligence: Too Narrowly Practiced", in Jennifer Sims and Burton Gerber, Vaults, Masks and Mirrors: Rediscovering U.S. Counterintelligence (Washington, DC: Georgetown U. Press), p. 175
[148] President (2008). "Executive Order 13470 -- Further Amendments to Executive Order 12333, United States Intelligence Activities" Federal Register 73:150, p. 45325; http://edocket.access.gpo.gov/2008/pdf/E8-17940.pdf
[149] President (2008). "Executive Order 12333 -- United States Intelligence Activities as amended by Executive Orders 13284 (2003), 13355 (2004), and 13470 (2008)", s. 3.5(a); http://www.ise.gov/sites/default/files/eo12333.pdf

> across the U.S. government. The effectiveness of security countermeasures in preventing penetration will be enhanced by intelligence concerning the current nature and scope of the adversarial intelligence threat. No single department or agency alone can ensure the integrity of the U.S. intelligence system and of our critical national assets and critical infrastructure.[150]

If the implication is that 'colleagues in security, acquisition, information assurance, and other relevant specialties' are not part of the counterintelligence community, this strategy is likely to neglect, rather than ensure, the integrity of the U.S. intelligence system.

CONCLUSIONS

In the course of our studies, we have discussed a wide-ranging and diverse threat to our nation's sensitive and classified information. This includes spies, but also leaks, hackers, self-anointed activists, screw-ups, psychopaths, and others. In many ways, the disappearance of our chief adversary, the Soviet Union, made protecting our secrets more difficult, not less so.

Defining counterintelligence in a full proactive scope is a necessary for formulating effective counterintelligence policy. EO 12333 no longer excludes personnel security, physical security, and information security, but nor does it explicitly include these activities as legitimate counterintelligence functions. They are indeed counterintelligence functions, and in the current environment even more important than counterespionage. While protection against spies remains an important task, it must not overwhelm the equally important task of protecting sensitive and classified information against any compromise -- whether from a foreign adversary or not.

In the information age, a robust, integrated, and comprehensive counterintelligence strategy must include personnel security, physical security, and information security among its functions. Regular and continuous auditing of classified information usage, such as provided by RTISFS, is essential. Thorough and careful vetting of cleared personnel, with regular review of their status, is also important. Good cyber-hygiene and well-protected information infrastructure are also crucial safeguards. Not only do these measures protect information from compromise in the first place, they also help create forensic knowledge necessary to identify, apprehend, and exploit persons who nonetheless seek to misuse that information, for whatever purpose.
Technical and policy means to ensure information is adequately safeguarded from misuse -- regardless of the motives of the person misusing it -- are the first, best defense against espionage.

---

[150] ONCIX (2009). "National Counterintelligence Strategy of the United States of America" (Washington, DC: Office of the National Counterintelligence Executive) http://www.ncix.gov/publications/policy/NatlCIStrategy2009.pdf; p. 2

RED TEAMING THE RTISFS CONCEPT

Stephen J. Lukasik

INTRODUCTION TO THE RED TEAMING METHODOLOGY USED HERE

The essence is to think like an attacker. He should be viewed as possessing a good grasp of current technology, not a superman or a magician, but also of greater capability of the *average* counterintelligence operator or law enforcement operator. Think like fast-track people (based on record, not political prowess.) Thinking like the adversary sharpens defensive thinking about feasibility, requirements, time scales, footprints, and indicators.

This enables one to understand the tradeoffs between technical sophistication of the attack, its time, cost and strategic objectives and thus aids in establishing priorities for the defense. It is important not to focus too narrowly on the single event, but to think at what might be called the *campaign* level. This is the reason for examining foreign intelligence collection styles

Thou output of a Red Team is, ideally, a detailed plan of execution. Time and resources have been inadequate in the instant project. However, the basic notion of a red team, to identify holes in what would be the defender's *modus operandi* under the RTISFS implementation as described in reports #1 and #2 of this contract, and the present report #3.

Reporting Red Team results to a sponsor is a delicate process, since it involves delivering criticisms that most operators of the systems under discussion will take as bureaucratic attacks on the dedication of their people. One can only quote Harry Truman: "If you can't stand the heat, get out of the kitchen."

Normally Red Team results have limited distribution since they do point up holes in system. In this case we choose to report them fully since failures of counterintelligence are substantiated in other parts of this report and the authors feel a review of holes are a contribution when facing a future when the penetration of information technology into all aspects of society are perhaps not completely appreciated by those who would protect "secrets."

Our methodology, absent thorough attack plans, is to start with the description of already defined and to point out holes we would exploit. A good result would be if the community dealing with such problems reports there is nothing new here. We hope that is the case here.

One point we make is that it is usual for a community of workers, in any field, to rely heavily on what has happened. These are obvious after the fact and statistics can be collected and analyzed for lessons learned. These lessons guide defender counteraction doctrine. This is sound but unimaginative. Attackers are imaginative because they derive different lessons from what has happened. The read the record for what will not work, and therefore they look for new approaches. Such approaches are hard to defend against because they have not happened, and are dismissed in favor of putting always scarce resources into the *real problem facing us*. When

attack technology is not changing rapidly, it is easier to think a few steps ahead, as was the case during the Cold War singular nuclear confrontation. All else was second-order.

But intelligence attacks are changing for three reasons. First they are based on information technology that *is* rapidly changing. And second, because the adversary has spread far beyond the KGB/FSB. This report reviews the way Russian intelligence may be changing and the way a new player, China, approaches questions of collection and analysis. The large increase in digital information adds a third element. There is a great deal of important information that all states release. Sometimes it is not a matter of policy but simply the way the world's population creates and distributes information technology.

In this final report, we attempt to capture these trends and put them into the context of future counterintelligence needs.

STIPULATIONS FOR READING THIS STUDY

First, that the RTISFS concept is technically valid.

Second, that RTISFS will work as described in this Phase I study, recognizing that what has been fleshed out has been limited by time and resources.

The point is that getting value from a Red Team analysis of a concept is not the time to raise technical objections. A Red Team study is not looking for technical flaws. It is looking for conceptual flaws. The understanding is that our concept could either be modified to eliminate the discovered exploit, or an operator could use it being fully aware of what I will not do. The premise is it is better to solve some problems than none.

Were the study to continue, we would develop the proposed software and test its efficacy in both laboratory and operational environments. The effectiveness of potential exploits would be subject to quantitative evaluation. But having chosen not to proceed is not to be interpreted as a lack of our confidence in our idea.

A further result of truncating this work is that a critical part of a complete Red Team study is missing. The value of a Red Team is to go on to study how to "fix" the conceptual holes. Since RTISFS software will not be developed, this step is moot.

EXPLOITABLE POINTS IN RTISFS

There are two principles followed here. First, trust no one regardless of their clearances. Clearances, while regarded as the gold standard (or perhaps the entry level – which is it?) in insuring against trust violations, by definition fail when trust violations occur. To rely on them is to enter into a closed circle of "honesty." To think about the problem one must step outside this closed circle and think like a counterintelligence officer. The second principle is that the more

people that must be party to a conspiracy the less chance it has of succeeding, and thus the less attractive it is to a trust violator.

There are fuzzy edges, however. Temporary clearances are often granted in critical circumstances, circumstances that could have been initiated precisely to obtain unauthorized access. One person "covering" for another out of a bond of loyalty represents an unwitting conspiracy. An insider may have a helper someplace, even if not in contact, since the two can be under the control of an outside coordinator. In addressing the problem of insiders, one must never forget the exceptions to whatever deduction one makes.

1. The "Top" Level

The "top" levels of government, or private organizations, that defines the sensitivity of information and the degree to which it must be protected, is *above* the chain of control it authorizes to act on its behalf. In its leadership role, the top level has the authority to release information if in its opinion, some greater national, or organizational, good is achieved. The top level may, however, be insensitive to the details professional analysts may derive for such top-level releases that can put lower-ranking people at risk or fails to protect fragile sources and methods. The top levels engage in unscripted foreign negotiations and these can also present opportunities for misjudgment, as when a person is manipulated to go further than they might have or through in an encounter with an adversary skilled in interrogation. There are no hard controls possible at top levels. Memoirs, biographies, and release of sealed information often point to critical points where this has happened. RTISFS can not play a hard role here since high-level information flows are unique to each leader's management style in terms of contacts and sources of private advice and support.

Nevertheless the top level still provides forensic opportunities for RTISFS to assist in negating hypotheses in investigating "unauthorized" releases and thus assisting in identifying violation paths for possible remediation. There can, even here, be overriding conditions, such as the personal security of a ranking person requiring ad hoc changes, perhaps under circumstances that can inhibit recording.

What all this says is the top is an messy place in which to maintain strict automated access rules and procedures. In physical terms, it is a porous membrane, one whose essential character must always be considered. That makes this level an ideal target for penetration since this is where the most useful information is to be found. [mention Roosevelt aide from Venona, the Cambridge five, etc.]

2. The RTISFS Management and Operational Level

This is the first level at which one has a defined structure to review for vulnerabilities. At this point it must be incomplete because, while the RTISFS operational management structure has

been outlined, many critical details of the scripting process that translates security policy to implementable rules, has not been designed.[151]

Manipulation of the protection afforded classes of information requires coordination between security officials representing a user's employer (assigning responsibilities to a user) and the "owner" of the information (typically the collector) to reach agreements on what information will carry what labels that are recognized by both parties. As a multiparty process it will be taken as unpromising for manipulation by an insider. The only concern is that if such responsibilities are taken as less important than other demands on the person responsible, the responsibilities will be delegated down to levels and people unknown, thus diffusing responsibility and affording opportunities for an inside manipulator.

The vulnerability that does emerge from the design of the RTISFS process is the Counterintelligence Controller (CIC.) As defined thus far, the CIC is the examiner of trigger event reports, the results of interrogatories, and such off-line analysis as he may call for. This is a major RTISFS vulnerability. Who watches the watcher? While it is beyond the scope of this report to recommend mitigation of this vulnerability, it is worth noting that there is a recursive ability in the RTISFS concept. There *can* be watchers of the watchers, an RTISFS system above the CIC level, that compares the behavior and performance of CICs. An RTISFS CIC will understand the capability of the underlying technology and will be a more difficult subject. One can have multiple CICs working in parallel and unknown to each other, with a super-CIC to use several input assessments on a suspect user.

It will be important to recognize that a CIC does not have to have the same access to the information he is protecting than does a user. This works both ways. The CIC will not be a suspect as an insider that way, although that could impede him in discharging his responsibilities. Even without access to the protected data, the CIC is an excellent target as an inside collaborator.

3. The RTISFS User Level

There are five ways for an insider to get around RTISFS, depending on the nature of the target information. Suppose the insider restricts himself to only information to which he has been granted access. But if the information to be taken out is the knowledge that there is nothing to take, that can be done without triggering RTISFS. Suppose the information to be taken is easy to remember. Then depending on how the foreign intelligence service operates, that may be all that is necessary. If the insider has a remarkable "photographic" memory, more complex documents can be removed the same way. If the insider is able to bring in a small camera, something increasingly feasible with information technology, a screen shot can be taken.

An even more powerful intrusion can occur if the insider is not supposed to take anything out. Instead his role is to put software *into* the system, software that can than make it possible for large numbers of outside collectors to do the actual removal. This points up the importance of

---

[151] The management process is described in our 3 August report: "Task 1 Forward-Looking and Backward-Looking Taxonomies." Stephen J. Lukasik, p.8-10 of 142; "Task 3 Design of RTISFS Functionality," Kenneth Hunter and Ted Russell, pg 31-50 of 142. [also Neal's description of three-factor reliability ref]

viewing the future CI problem *ab initio* and not simply "cyberize" traditional intelligence collection.

4. "Universal Access" Insiders

There are classes of insiders having nothing to do with information handling that must have access to all spaces: support staff assuring an area is safe for occupancy, that users are not harmed by high risk materials and activities, first responders to emergencies, etc. They can, of course, be cleared but the premise is that this can not be relied upon. They are often cleared less rigorously than information workers and are often invisible/nuisances. But their access gives them great capability to find out what is going on in a closed facility. For example, in WW II Soviet agent David Greenglass provided details of nuclear weapon trigger mechanisms based on his access as a health physicist at Oak Ridge.

Such people can be seen as having no access to critical details, but they can use their senses and overheard remarks based on their physical access to spaces to assemble a larger picture. Today, their access to an organization's information system for filing their reports and be used penetrate further, to input software to aid an outsider, or to engage in social engineering to further access.

5. Psychopaths

Noted in our 3 Aug report, and in this report, psychopaths pose particular problems for RTISFS since their mental processes differ from those of people that make up the normal training set for RTISFS. Such people can lead lives seen as "normal," and their recruitment my an attacker could, therefore, be quite promising. A recruiter, in selecting such people for recruitment presumably would compensate for their in condition in handling them, though they could provide an interesting special case for turning.

6. Departing employees

Insiders who leave, either as a result of recruitment by an attacker for this purpose, or through their administrative actions, pose an important subset of insiders. We have seen a few cases where such people "stockpile" information for possible future use. The day after a person leaves they are as well-informed as the day before. They also, could b useful in emplacing leave-behind devices or software. Their user behavior in this period can normally be expected to change so anomalies can be attributed to this life change. This will have to be addressed in the design of RTISFS.

7. Travelers

The ubiquity of mobile devices and use of new information and communication technology offers considerable latitude for attackers to modify traveler devices. Analysis of this universal situation for defense of the network is larger than the insider problem understood as a witting insider. High risk behavior will be increasingly common independently of a user decision to violate trust. This broadens requirements on online analysis of users since, like psychopaths, lessons learned from the past become increasingly irrelevant. Like the increasing fraction of

information-should-be-free users, unwitting users may far outweigh witting inside penetration agents.

CONCLUSIONS

There are five conclusions to be drawn from this discussion, as well as consideration of CI futures.

1. There are two uncertainty principles in operation, analogous to the Heisenberg analysis of wave-particle duality

    For the defender: Δ(security) Δ(utility) $\geq$ some constant

    For the penetrator: Δ(precision) Δ(timeliness) $\geq$ some constant

2. Watching the watcher must be considered in the design of RTISFS. It has not been in the work reported under the Phase I contract.

3. While an RTISFS system, as conceived as a stovepipe, can be tightly implemented, the top and the bottom levels are particularly prone to leakage because they are difficult to manage.

4. The dynamic nature of the interaction between attacker and defender is unavoidable, and thus it must be built into defensive measures. Static defenses should be avoided; the scalability of defenses is a requirement; and technology cuts both ways.

5. Learning lessons is important. Relying on those lessons is treacherous.

TASK 3 – RTISFS FUNCTIONALITY AND SOFTWARE DEVELOPMENT PLAN

Frank J. Sauer

As we have undertaken research to sharpen our proposed idea for the ADAMS SBIR program, we have arrived at two positions quite different from those with which we started. First we see important ideas for more fundamental research on the mental processes that lie behind perceptions of truth and the importance of truth as a motivator of behavior. The second is we see how the details of software mechanisms for accomplishing the goal of the ADAMS program are relatively unimportant compared to the former. As we came to this recognition we reduced our efforts on Task 3 and transferred research resources to the examination of the more fundamental issues addressed in this report.

As a result, this section present a concise discussion of the work plan that could be used to move from the totality of the three RTISFS SBIR reports to a project to develop RTISFS as operational software, and includes: a descriptive summary of RTISFS software; the RTISFS functionality requirements; recommended software development methods to be used throughout the project; and the description of the necessary specific project plan tasks (i.e., task descriptions, milestones, deliverables, and technology readiness levels).

SOFTWARE DESCRIPTION SUMMARY

The RTISFS software suite is a user-configurable data-centric software tool designed to map data-use anomalies, with the main purpose and focus on identifying activities and patterns typically related to a potential malicious insider. As data is used – whether it is read, copied, deleted, edited, cut and pasted, renamed, etc. – there is a trail of data activity that will be generated, monitored and recorded. When this trail of activity is analyzed, patterns of data-use become evident. Further, these data patterns are linked to origination source, whether from one or multiple clients or sources, to provide attribution to a person or persons of potential interest.

The suite's goal is to identify data patterns of interest that go outside the normal data use. This can be achieved once RTISFS is installed; a "data-use" baseline is established with data of interest, profiles are developed and installed; and a configuration element is available for unique data-use identification.

Once an abnormal data pattern is identified, a systematic flag is generated to alert that an anomaly occurred. At this point the Counter-intelligence controller (CIC) may generate a task to RTISFS to generate a report specific to the flagged activity. Based on how the report is configured, it will return information and links relative to the data in question. This report will not only identify the data in question and its unusual nature, but will also link back to any source or sources that are generating the anomaly (but not the data itself, just the metadata).

As a tool, RTISFS is not a complete solution to any matter associated with data in a network. Although powerful and dynamic in construct, it is only as effective as it is configured.

Additionally, the more data it has access to, the more effective the data patterns for identifying anomalies.

RTISFS FUNCTIONALITY REQUIREMENTS[152]

The recommended RTISFS initial implementation platforms are Windows® 7 client in Intel or AMD-based platforms and Windows® Server 2008 R2 in Intel-based platforms.

RTISFS does *not* change the actual enterprise classification of the data itself. Therefore, for RTISFS purposes *only,* all data imported from an external sources (e.g., mainframes and servers not within the defined RTISFS-enabled network) is treated by RTISFS at the user's access level as configured by the RTISFS administrator. Only by setting such external data to the highest level of the recipient's access restriction within RTISFS, can we protect it in RTISFS terms.

There are three key concepts at the foundation of this RTISFS functionality: data-centric, interrogatories, and controlled outcome responses.

Data-centric Concept. This is the model (or construct) that puts data and their use at the center of any and all operations taken; whereas, all processes and procedures are external and supporting of the data, or leveraging the data, or both. Since all computers function in the data realm, it is the process of interfacing and functionally using the computer/network in its *native* operational environment. In the case of RTISFS, the data are mapped contingent with the data access/use. The benefits of this are:

Simplification of problem space. By addressing anomalies *in the data*, the points of consideration are drastically reduced and more clearly identified. By contrast, in observing *the users and their actions*, the problem space becomes too big and too complex. When this happens, the relevance becomes a game of guesswork or extrapolations. And the correlation to behavior and motivation is lost in the model, especially when potentially multiple people are involved.

Relevance. Humans are identified by their behaviors. Alternately in the data-centric realm, behaviors (for RTISFS) are defined only as "touching data" or in changes to normal patterns.

Correlation. Only by using data-centric methods and profiling is there any chance of *identifying collusion* through use of data on the network. A good example is if the focus is on *two* separate users; their individual data-use activities in and of themselves may not draw attention as unusual. Yet there is collusion between them. In a data-centric focus, the data-use pattern would likely reflect the colluded data pattern, and then the flagged data anomaly would point back to both users.

Interrogatories Concept. Since people are not consistent with the way they carry out actions, especially on a computer, the interrogatives component is installed as a means to fill in the gray

---

[152] This describes the RTISFS foundation document (i.e., design criteria/technical specification). A more detailed description is available as Task 3 in the RTISFS Phase I Interim Report, dated August 4, 2011, and available through the DARPA ADAMS Program Manager.

areas that data anomalies alone would not identify. The interrogatives can be automatically or manually generated. The interrogative presents an interactive component to the user in order to evoke a response for action taken on the client. Further, the component allows a real-time manual interactivity to allow the CI Controller the flexibility of evoking immediate responses by the user for potentially questionable actions taken.

The interrogatory works in conjunction with the profile. When monitoring, there are two data matters with which to deal: profiles to look for the norms in the data and the interrogatories to determine and help sort through the non-norms.

Controlled Outcome Responses Concept. The RTISFS Alerting component puts RTISFS into the Response mode. In this mode, the matter of reacting to the alert is brought to the forefront, providing multiple venues in addressing the alert.

In this Response mode, RTISFS allows the CIC to become interactive with the potential insider by directly sending an interrogatory to the user. At this point the scripting or insertion of extra interrogatories plays out—CIC actions are generated. The reason Controlled Outcome is different is because RTISFS is real-time interacting with the action of the computer to resolve an issue that the CIC is investigating: "Is this person of interest really a 'person of interest' ?"; "How do I need to find out what is going on?"; "Do I need to change representation?"

RTISFS can change the system representation to focus the RTISFS Alerting component on a specific "Person Of Interest" (POI). It is likely that every POI will need to be handled individually—meaning there is no one checklist as to how to handle all POIs. By providing the scripting mechanism, we are providing this component the means so it may be *customized* to any response in the way the organization/CIC sees fit to investigate each individual POI.

At the point of characterizing an individual as a POI, it is assumed that legal involvement must be working hand-in-hand to ensure all interactions are carried out legally. If there is need for a "canned" standard response, it is necessary to ensure it is legally approved for *all* situations where it is to be used. If an interrogatory is to deviate from pre-scripted approved interactions, legal assistance must be involved, and documents approval for the deviation is likely to be required.

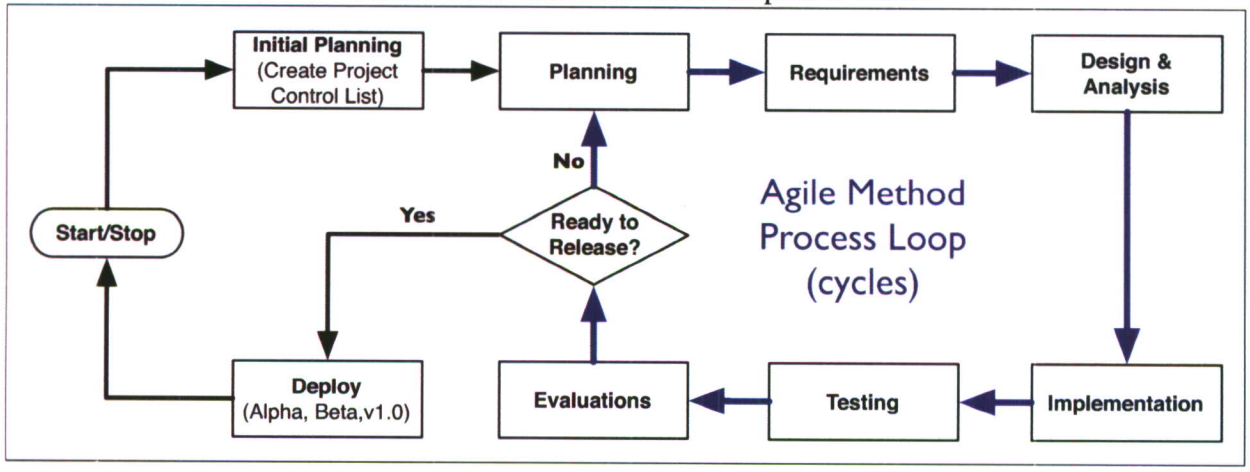
THE RECOMMENDED METHODS TO BE USED IN THE SOFTWARE DEVELOPMENT

It is recommended that the RTISFS software development rely on an Iterative and Incremental Development Model using a modified agile methodology, which is central to the software development process. Through this cyclic process the expectation is that each developed component will be functional before moving on to the next cycle. Also, this agile method will be central to keeping all the developers fully aware of each area of development throughout each of the forecast cycles.

The idea behind the agile methodology for developing a software system through repeated cycles (Iterative) and in smaller portions at a time (Incremental), is that it allows the software developers to take advantage of what was learned during development of earlier parts or versions of the system. Learning comes from both the development and use of the system, where possible

key steps in the process start with a simple implementation of a subset of the software requirements and iteratively enhance the evolving versions until the full system is implemented. Each iteration results in design modifications that are made and refined and new functional capabilities are added, as illustrated in Figure 10 below.

Figure 10
Software development process
Iterative and incremental development model



The procedures for the Iterative and Incremental Development Model consists of:

The Initialization Step. This step creates a base version of the system on paper. The goal for this initial implementation is to create a product to which the prospective user can identify the technical functionality of the system. It will offer a *sampling* of the key aspects of the problem and provide a solution that is simple enough to understand and implement.

The Project Control List. To guide the iteration process from start to finish, a *Project Control List* is created that contains a record of all tasks that need to be performed/developed. Initially, it is the development plan with software development boilerplates for each RTISFS component. Through each iteration cycle, it will include items such as new features to be implemented and areas of redesign of the existing solution. The Project Control List is constantly revised as a result of the analysis phase.

The Iteration Step. Each iteration involves the redesign and implementation of a task from the Project Control List, and the analysis of the current version of the system. The goal for the design and implementation of any iteration is to be simple, straightforward, and modular, supporting redesign at that stage or as a task added to the Project Control List. The Project Control List is modified in light of the analysis results.

Understanding this development model and agile methodology, it is suggested there be three independent development streams that reconvene and integrate at the end of every agile cycle. As a result this approach includes an increased frequency of the scheduled intervals, and more demanding methods of review, testing, and updates.

We also suggest a sub-cycle structure in this model. There are 3 distinct phases:

(a) design review and design update and cycle specific declinations;
(b) unit development and unit testing, which is a continual cycle and iteration in that phase until every aspect described in (a) is complete;
(c) component integration. Component is comprised of one or more developed units. Integration, comp. testing, defect id and resolution.

Combined, with these changes the development will maintain a more accurate picture of the whole development project through updating the Project Control List frequently, and by exiting each cycle with *fully functional* components that are ready for integration with other components upon development. This process will increase the likelihood of minimal rework time and troubleshooting when objects and components are combined to create complete components and also the complete prototype, Alpha version, Beta version, and v1.0 release product.

Finally, this modified Agile methodology will promote a close cooperation among the developers, each of whom will be fully aware of the overall state of progress as it relates to the overall development strategy. We have found this process will foster a working environment that leads to more efficient code as well as a much smoother integration of components because the ongoing interaction fosters greater understanding among and between all of the development members.

RTISFS SOFTWARE DEVELOPMENT TASKS

At the point of starting the development process, the effort would be at the TRL 3 level.[153] That is, enough analysis has been done to convince the developers that the software can be made to work. Algorithms needed to perform the major functions have been worked out, and some exploratory coding has been done to see whether some of the more speculative ideas can be gotten to run on a laboratory computer. There is no attempt to integrate functions and/or databases. I/O is done manually, or in a "brute force" manner. The developers would have also researched if they can reuse or adapt existing software to the project needs.

PRE-CYCLE TASK: LOW-LEVEL SOFTWARE DESIGN.

Description: There are two parts to this task. The first part is familiarization for the development team with the RTISFS project and the existing designs. The second part is the all hands design work on the low-level design. While the succeeding cycles will have each developer working on his own separate stream, this effort will be done in common, to insure that everyone is on the same page and that all the pieces will fit together.

The common effort is on the low-level designs that will build the source code, framework, and templates (stubs) utilized in the development process, such as the blank Object-oriented Classes of the system. At the end of this task, all the

[153] All the TRL comments are adapted from TRL definitions specifically for software presented in the TRL Calculator v2.1 and available on Internet at URL: http://www.acq.osd.mil/jctd/TRL/Documentation%202_2.xls

design specifics of how the various components of RTISFS software will fit together, at the lowest level, will be specified.

Duration: 2.0 months

Milestone: +2.0 months from start of the project; Initial Low-level design complete.

Deliverables: None.

TRL: working toward TRL 4

CYCLE 1 TASK: MONITORING

Description: Beginning with this cycle, there will be three (3) development streams; an individual developer is completely responsible for a single stream. Overall, this will create the components for establishing RTISFS configurations for implementing organizations, and all the setting up of RTISFS data profiles[154].

The first development stream is the Administration Tool. This creates the ability to designate specific data files as sensitive, specialized or neither. This also covers the definitions, the groupings and how to set up the profiles (which is the container that holds the thresholds that are the basis of the "sensor" triggering in RTISFS).

The second development stream is the Client–Server setup and the communication tool between them. This includes the data definitions, how they are stored and protected, and the version control process for the RTISFS scripting schema.[155]

The third development stream is the Monitoring of profiles. On the client side, this is determining what data streams to monitor and what the data construct is for that profile monitoring. The internal data sources within RTISFS will be protected by encryption and strict access controls at this point, and log entries cannot be altered. The data classification will be assigned inside the client system, as defined within the Admin tool. Internal communication within the RTISFS system will be in encrypted XML format to allow scalability.

---

[154] An RTISFS profile is a programmatic term that describes a configurable record that holds references of activities being performed on a set of data. In referring to those activities, thresholds are developed for those activity patterns that are maintained in a record documenting the thresholds. Crossing the profile thresholds is what triggers an alert to the CI controller. The organization may provide additional information to the profile templates to expand the construct of profiles specific to their data monitoring needs. Details are available in the Task 3 section of the RTISFS Phase I interim Report of August 4, 2011, available through the DARPA ADAMS Program Manager.

[155] Note: this version control is *not* referring to the versioning control of the development process. It is a component of the RTISFS software.

Duration: 5.5 months; 0.5 month; 4.0 month; 1.0 month[156]

Milestone: +7.5 months from start of the project.

Deliverables: "Show and Tell" of the Administration Tool. It will show what can be configured, and a task screen showing that there is a client attached. Client operations will not be observable at this point.

TRL: TRL 4 — This level of technology development is primarily concerned with the coding of individual modules and/or functions. The developers do some ad hoc integration, but it's still primarily force-feeding the output of one module into another with little regard for the final interface characteristics. The developers know enough about the software project to perform initial estimates of software size for use in program risk management and cost estimation.

CYCLE 2 TASK: REPORTING

Description: Cycle 2 covers the preset Interrogatory windows, the Triggers, the Counter-Intelligence Controller (CIC) interface for modifying Profiles, and Reporting.

The CIC will not be able to respond within RITSFS, but the data generated can be used to generate CIC responses manually. The monitoring is being done automatically by the RTISFS software.

The Data classification is the only interrogatory introduced at this point. When they create raw data (create a file from scratch), they will have to classify it manually by the criteria set up by the Admin Tool. A trigger is sent to the CIC if the classification varies from the rules set up by the Administrator. It accepts their input, but the trigger will still go back to the CIC.

Starting with this Cycle 2, at the end of the Cycle will be a "pre-Cycle" that will be broken into a week of testing, a week of rework, another week of testing, and another week of rework. This is integration testing of all the completed components of all the developers. It is in addition to each developer's required "internal" testing of each of his component efforts before he can bring that component to this integration testing at the end of each cycle.

Duration: 5.0 months: 1.0; 3.0; 1.0

Milestone: +12.5 months from start of the project.

Deliverables: "Show and Tell" of the CI Analyst Tool. This will show the results of the client operations.

TRL: working toward TRL 5

---

[156] The breakout of the cycle total duration is for the 3 sub-cycles described in 3.2 above and for sub-cycles a., b., and c, respectively.

CYCLE 3: INTERROGATORIES

Description: CI has the ability to create the standard interrogatories that are generated in response to a trigger (and also random). RTISFS will keep a record of those responses and provide that information back to the CI on demand.The user's answer to an interrogatory is the action as well as any text that the user provides.

This longer than normal sub-cycle A is to allow the extra time for the scripting of primitives that are the basis of the interrogatories and the customized responses of cycle 4.

This development continues into Cycle 4, after the pre-cycle testing.

Duration: 5.0 months: 2.0; 2.0; 1.0

Milestone: +17.5 months from start of the project.

Deliverables: None.

TRL: TRL 5 — The developers still working in a laboratory environment, although they may be working with a processor that is representative of the target system application. Software architecture is well defined, and they can perform laboratory level integration of software modules and functions with lab data. A configuration management scheme and test protocol are defined and documented.

CYCLE 4: CUSTOMIZED RESPONSE

Description: Integration into CI tool to create targeted scripted actions for individuals, includes what triggers the action,what action should be taken (scripting macro) and the associated reporting requirements. This cycle is the first that can trigger to a specified user; one that is based on data and the user in combination. No customized individual action will be created without the CI setting it up.

The developer are also opening up their data repository through API's for read-only purposes. Use of that integration is for 3rd party data analysis tools/ integration into other Security Operations Center systems.

It is not an alpha. It is pre-alpha

Duration 4.5 months; 1.0; 3.0; 0.5

Milestone: +22.0 months from start of the project.

Deliverables: "Show and Tell"of enhanced CI Tool and 3rd Party integration

TRL: working toward TRL 6

FULL SYSTEM INTEGRATION AND TESTING

Description: This closes out the Cycle 4 effort with final testing of the Alpha RTISFS. It is putting every aspect together and testing for problems. This is where the performance metrics are measured and problems identified for modifications.

This is *not* tested in an operational sense. It is tested in a limited laboratory test at the development offices.

Duration: 2.0 months (integration of everything into one package and its testing)

Milestone: +24.0 months from start of the project

Deliverables: Demonstration of RTISFS *Alpha*. Not Deployable

TRL: TRL 6 — This is the first attempt to subject the software as a system to a realistic albeit simulated operational environment. The developers can expect numerous bug fixes and upgrades as deficiencies are discovered. Software is at the prototype level. Releases are "Alpha" versions and are under configuration control.

CYCLE 5: ALPHA EVALUATION

Description: This period will be focused on installing RTISFS in an identified operational environment for closed testing. As RTISFS is configured in the operational environment, project development personnel will be working side-by-side with the users to configure, develop scripts, and put RTISFS online.

Throughout the cycle, project technicians will be observing RTISFS functionality, documenting areas of concern, working with users to receive operator feedback, and ensuring data flow continues unimpeded.

Not just defects, but change requests from the user based on the process experiences, are documented. These changes will include changes in required functionality, such as changes to ease CI in configuration settings. All issues, and additional needs will be documented for repair, modification, and addition in Cycle 6.

Developers will also be working on documentation.

Duration: 6.0 months (no sub-cycles)

Milestone: +30.0 months from start of the project.

Deliverables: None.

TRL: working toward TRL 7

CYCLE 6: BETA DEVELOPMENT

Description: The alpha is uninstalled at the end of cycle 5, as the changes needed are likely to be significant, when the process has been observed. All documented problems and needs from cycle 5 will be addressed during this cycle back at the development center. An extensive testing will wrap up the cycle, and once completed successfully, RTISFS will be a solid BETA product.

Duration: 6.0 months: 1.0; 4.0; 1.0

Milestone: +36 months from start of the project.

Deliverables: RTISFS Beta software. (limited open test in a limited enclave)

TRL: TRL 7 — The developers are still working with a software system prototype; any software releases will be configuration controlled "Beta" versions subject to operational (field) conditions. The verification step of VV&A is completed, demonstrating that the software prototype meets the established requirements as documented in the software system specifications.

CYCLE 7: BETA EVALUATION

Description: RTISFS BETA will be installed in several operational environments, using the existing system settings (profiles and schema settings). A new baseline will be created.

RTISFS will be left running in the operational environments under the purview of the users. Project personnel will not be with the users during this cycle, however, communications with the users will continue in order to gather any further information as to problems or needs. Feedback can include remote connectivity to a beta testing facility or travel to that facility, as deemed appropriate. Also, the developers will fix anything that hinders use of RTISFS during this cycle.

Developers will only observe user system use and interaction, documenting any areas of concern, if any, and assisting the users when asked.

Developers will use slack time to further develop documentation, technical and operational.

Development of a training program,and training beyond what has already been provided will be carried out as requested.

Duration: 6.0 months

Milestone: +42 months from start of the project.

Deliverables: None.

TRL: working to TRL 8

Cycle 8: Release Candidate

Description: All final documented matters of concern or need will be repaired and updated during this cycle. There will be a final rigorous testing that will result in the final product.

Project personnel will install the final product as requested by the Government and a new baseline will be created in the installed network.

Training will be conducted on how to install and integrate RTISFS in other networks, setting them up, and putting them online with the required personnel to manage and utilize RTISFS.

Technical and Operational Documentation will be delivered.

This concludes the development and implementation of the initial releasable version (1.0) of the RTISFS tool.

Duration: 6.0 months: 1.0; 3.0; 2.0

Milestone: +48 months from start of the project.

Deliverables: Version 1.0 of RTISFS software.

TRL: TRL 8 — The software is ready to be installed in an operational environment.